American Journal of Computational Linguistics *Microfiche 57*

PHONOLOGICAL RULES FOR A TEXT-TO-SPEECH SYSTEM

SHARON HUNNICUTT

NATURAL LANGUAGE PROCESSING GROUP

RESEARCH LABORATORY OF ELECTRONICS

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

CAMBRIDGE 02139

This work was supported originally by the Joint Services Electronics Program, Contract DAAB07-71-C-0300, and more recently by the National Science Foundation (Grant EPP74-12653)

Summary

The phonological rules discussed in this paper are part of a system which has been under development at M.I.T. to convert unrestricted text to speech. The system utilizes a morph lexicon and a vocal tract model. Although most of the linguistic analysis is done by decomposing words into their constituent morphemes, such a system is not sufficient for unrestricted text. In order to attain the competence of a comprehensive system, it was necessary to develop a scheme for dealing with unrecognizable words. This is called the "letter-to-sound" system.

When a decomposition fails, that is, when a word cannot be decomposed into its constituent morphs or when it is too infrequent in the English language to be included in the morph lexicon, the "letter-to-sound" system is invoked. The letter string which it receives is converted into a stressed phoneme string using two sets of ordered phonological rules. The first set to be applied converts letters to phonemes, first stripping affixes, then converting consonants and finally converting vowels and affixes. The second set applies an ordered set of rules which determine the stress contour of the phoneme string.

These rules were developed by a process of extensive statistical analysis of English words. The form of the rules reflects the fact that pronunciation of vowels and vowel digraphs, consonants and consonant clusters, and prefixes and suffixes is highly dependent upon context. The method of ordering rules allows converted strings which are highly dependable to be used as context for those requiring a more complex framework. Detailed studies of allowable suffix combinations and the effect of suffixation on stress and vowel quality have also provided for more reliable results.

Table of Contents

Figures

Examples

Letter-to-Sound

Lexical Stress Placement

A Complete Example

I. The Text-to-Speech System

A system to convert unrestricted printed text to speech has been under development for several years at M.I.T.[1,2,3] The approach has been to model the process employed by a native speaker of English when reading aloud. In order to develop correct computational algorithms for the pronunciation of English words, it has been necessary to reflect the basic nature of linguistic processes. Consequently, considerable emphasis has been placed on the development of morphological and phonological analysis, stress patterns, parsing systems and prosodic correlates.

It is possible, using the current system, to convert any English word or string of words in a textual representation to intelligible speech. In order to effect this conversion, a number of subsystems are utilized including morphological analysis, letter-to-sound rules, stress rules and phonemic speech synthesis. Prosodic studies are now in progress; experimental parameters for f_0 contours and timing patterns will soon be included in the system.

The letter string which represents a word is usually converted to a phoneme string in preparation for speech synthesis by a process of morphological analysis. A morph lexicon containing approximately 11,000 entries has been developed and is used in conjunction with a morph decomposition algorithm. Included in the lexicon are two major classes of morphs. One class is composed of roots such as "trust" and "snow," i.e., words which can occur alone, and bound roots such as -ceive perceive, receive, conceive), rot- (rotary, rotate, rotor) and -miss- (dismiss, missive, permissiveness). Affixes make up the second class and may be attached either to roots or to

bound morphemes. Accompanying each lexical entry is its phonemic representation, its morph class and its part(s) of speech.

Algorithmic decomposition of letter-strings models the procedure used by a native speaker when confronted by a word which he does not immediately recognize or has previously not encountered. If the word is not immediately recognizable, i.e., if it is not in the "lexicon" of the native speaker in its entirety, an attempt will be made to break it apart into its constituent morphs. Such a process is probably used when one reads a word such as "antidisestablishmentarianism," earthrise" or "cranapple" for the first time

The algorithm also models the ability to recognize mutations such as the dropping of a final silent [e]$^{-4}$ (observe - observance), the doubling of a final consonant (red-reddest) and the substitution of [i] for final [y] preceding vocalic suffixes (glory - glorious). Morphophonemic rules are also included, modeling the ability to give a correct pronunciation for any plural (horses, cats, dogs) or past tense (quieted, hushed, whispered), and to palatalize in appropriate contexts (sculpt - sculpture, confuse - confusion).

Another feature of the algorithm is a set of selectional rules which, although very simple in form, choose the correct morphemic analysis from all possibilities in a large number of cases. A standard form for sequences of morphemes is compared with each possibility, and rules describing preferred composition are used in a pairwise comparison leading to an acceptable result. The word "formally" provides an example in which the only rule needed is one stating that a root followed by a suffix is preferable to two concatenated roots. Possible decompositionsof the word "formally" are

detected as follows(R represents "root" and S, "suffix"):

form (R) + al (S) + ly (S)	form (R) + ally (R)
for (R) + mall (R) + y (S)	form (R) + all (R) + y (S)

It is clear that the correct decomposition is the only candidate having the form of a single root followed by suffixes.

When a complete morphemic analysis fails, that is, when a word encountered by a native speaker cannot be separated into its constituent morphemes, or when it is so infrequent in the English language that he has previously not encountered it, a "letter-to-sound" system is invoked, i.e., an attempt is made to sound out the word letter by letter. The competence of a native speaker which allows him to perform this conversion is based on correspondences between letters and sounds in English which have been internalized through experience. (A native speaker will also apply the same correspondences to a foreign word from any language of which he has no knowledge.) A scheme to model this process must be made available, sequenced after the decomposition algorithm, in order to be able to convert unrestricted text to speech. The text-to-speech system includes a phonological model having a two-phase structure: the first phase is a set of rules which converts letters to phonemes, and the second is an algorithm for placement of stress on the converted phonemes.

This system has been implemented in BCPL on a DEC PDP-9 and on a PDP-10 in MAC LISP.

One might ask why, with a set of letter-to-sound rules available, it is necessary to have a lexicon. There are three reasons; all are restrictions which must be imposed on any viable letter-to-sound system. First, it has

been observed that high-frequency words perhaps because of extensive use, do not always follow letter-to-sound rules. For example, the only instance in which a final [f] is pronounced as /v/ is in the word "of " The letter [w] following a consonant is generally pronounced like the [w] in "sweet" or "twill," but in the word "two" it is not pronounced at all. A study[5] of the 200 most frequent words in English according to the Brown Corpus[6] was made to determine their regularity of pronunciation. It was found that although the regular case is that a final [e] preceded by a single consonant (other than [r]) lengthens the preceding vowel, four of the 200 words, i.e., "have" (compare "behave," "shave"), "one," "some" and "come" (compare "lone," "ode") are exceptions. The case of initial [th] is even more irregular among high-frequency words. In most English words, initial [th] is unvoiced as in "thistle," "thin," "thesis," However, twelve of the 200 words began with voiced [th].

Secondly, it must be recognized that the letter-to-sound rules which operate within a morpheme do not necessarily apply across morph boundaries. In particular, the pronunciation of compounds requires a lexicon. Such words as "hothouse" or "potherb" might otherwise appear to contain the consonant cluster [th], and the morph-final silent [e] in "houseboat" would certainly not be silent if the word were not recognized as a compound. The application of letter-to-sound rules must therefore be restricted to words containing no more than a single root.

Thirdly, foreign words which retain their original pronunciation must be lexical entries. The entries may be made in the same way a native speaker of English would add a foreign word to his vocabulary, i.e., by pronouncing it

as if it were an English word (using English letter-to-sound rules) until informed of its correct pronunciation, and then placing it in his mental lexicon.

It is apparent, then, that both morphological and phonological systems are necessary and that together, in sequence, they can provide a phonemic representation for any English word presented for conversion to speech.

Although, at present, there is no interaction between the decomposition and letter-to-phoneme algorithms, a more highly efficient system could be developed in the future. The size of the lexicon could be reduced, for example, by application of stress placement rules to the output of the decomposition algorithm or by omitting unnecessary phonemic representations

II. Letter-to-Sound

The conversion of a letter string to a phoneme string in the letter-to-sound program proceeds in three stages. In the first stage, prefixes and suffixes are detected, (cf. Figure 1) Such affixes appear in the list of phonological rules. Each is classified according to (1) its possible parts of speech, (2) the possible parts of speech of a suffix preceding it, (3) its restriction or lack of restriction to word-final position and (4) its ability to change a preceding [y] to [i] or to cause the omission of a preceding [e]. Prefixes are given no further specification.

Detection of suffixes proceeds in a right-to-left, longest-match-first fashion. When no additional suffixes can be detected, or when a possible suffix is judged as syntactically incompatible with its right-adjacent suffix by a parts-of-speech test using classifications (1) and (2) above, the process is terminated. Finally, prefixes are detected left-to-right, also by longest match first. If at any time the removal of an affix would leave no consonant or no vowel in the remainder of the word, the affix is not removed.

Example: dictatorship

dict + ate + or + ship		possible suffix analysis
ship:	(a) nominal suffix	
	(b) follows nominal suffix	
or:	(a) nominal suffix	
	(b) follows verbal suffix	
ate:	(a) verbal, nominal and adjectival	
	(b) follows verbal, nominal and adjectival suffixes	
dict + ate + or + ship		parts of speech are compatible; analysis accepted.

Example: passing

pa + s + s + ing — possible suffix analysis

- ing: (a) nominal and verbal suffix
- ing: (b) follows nominal or verbal suffix.
- s: (a) nominal and verbal suffix
- s: (b) follows nominal and verbal suffixes
- s: (c) appears only in word-final position — unacceptable analysis

pass + ing — correct analysis

Example: finishing

fin + ish + ing — possible suffix analysis

- ing: (a) nominal and verbal suffix
- ing: (b) follows nominal or verbal suffix
- ish: (a) adjectival — parts of speech not compatible
- ish: (b) follows nominal or adjectival suffix

finish + ing — correct analysis: root functions as verb with verbal ending, ish

The domain of application of the second stage rules excludes any previously recognized affixes and is assumed to be a single morpheme. This stage is intended primarily for consonant rules and proceeds from the left of the string to the right. Extending the domain to the whole letter string once again for the third stage, a phonemic representation is given to affixes and to vowels and vowel digraphs, (cf. Figure 1).

Phonemic representations are produced by a set of ordered rules which convert a letter string to a phoneme string in a given context. Both left and right contexts are permitted in the expression of a rule, and may contain variables as well as letters or phonemes Any one context may be composed of either letters and letter variables or of phonemes and phoneme variables. Combination of these possibilities for both left and right contexts allows for four possible context types. One type of rule, for example, makes it possible to convert a particular letter string to a phoneme string only if the left context is a specified phoneme string and the right context is a specified letter string.

The method of ordering rules allows converted strings which are highly dependable to be used as context for those requiring a more complex framework. Because the pronunciation of consonants is least dependent upon context, phonological rules for consonants are applied first, i.e., in the second stage. Rules for vowels and affixes, requiring more specification of environment, are applied in the third and final stage. With the benefit of a previously converted consonant framework and the option of including as context any phoneme to the left of a string under consideration, the task of converting vowels and affixes is simplified.

DENOMINATIONS	Input	
DENOMIN + ATE + ION + S	Stage 1:	(a) recognition and isolation of suffixes
DE = NOMIN + ATE + ION +S		(b) recognition and isolation of prefixes
-- = n-m-n + --- + --- + -	Stage 2:	conversion of consonants in root
dɪ = n-m-n + --- + --- + -	Stage 3:	(a) conversion of prefix
dɪ = namɪn + --- + --- + -		(b) conversion of vowels in root
dɪ = namɪn + eʃ + ən + z		(c) conversion of suffixes
d_{2}ɪ = nam_{1}ɪn + eʃ + ən + z	Result of Stress Placement Rules	

All phonemes are given in IPA symbols. A dash (-) serves as a place-holder for a letter which has not yet been converted; an equals sign (=) follows each prefix; a plus (+) precedes each suffix The result of stress placement rules is also given.

Figure 1

Application of Letter-to-Sound Rules

Within the two sets of rules for conversion of consonants and vowels, ordering proceeds from longer strings to shorter strings and, for each string, from specific context to general context. The rule for pronunciation of [cch], then, appears before the rules for [cc] and [ch], each of which is ordered before rules for [c] and [h]. Procedures for the recognition of prefixes and suffixes also require an ordering: the prefixes [com] and [con] must be ordered before [co]; any suffix ending with the letter [s] must be recognized before the suffix consisting of that letter only.

As an example of ordering rules for a particular string, consider the vowel [a] and assume that it is followed by the letter [r]. This [a] may be pronounced like the [a] in "warp," "lariat" or "carp" depending upon specification of further context. It is pronounced like the [a] in "carp" if it is followed by [r] and another consonant (other than [r]) and if it is preceded by any consonant phoneme except /w/ (note "quarter," "wharf") Consequently, a rule for [a] in the context of being preceded by the phoneme /w/ and followed by the sequence [rC][7] is placed in the set of rules. Specification of a left context in the rule for the [a] in "carp" is subsequently unnecessary. If the [a] is preceded by a /w/, this rule will never be reached; if preceded by a vowel, a rule for vowel digraphs will already have applied. Using this method, rules may be stated simply and without redundancy.

Development of the set of phonological rules was begun by informal inspection and reference to published works, e.g., Venezky.[8] By a process of extensive statistical analysis, other rules were added and ordered

appropriately. The principal source of words was the Merriam Webster Pocket Dictionary.[9] A computer print-out was generated in which all words containing each letter and each specified cluster of letters were isolated. Within each category, words were sorted alphabetically according to the right-hand context of the letter(s) under consideration. In addition, Walker's rhyming dictionary[10] was used to determine pronunciation of suffixes and the effect of suffixation on preceding phonemes. Words from the Brown Corpus, the Heritage English Dictionary[11] and Stedman's Medical Dictionary[12] have been used in testing procedures.

Examples of Rule Application

In this section, a number of words will be analyzed according to the phonological rule program. Intermediate output, i.e., the results of the first and second stages, will be provided for each word, and the rules which have been applied to produce this output will be discussed. Generalizations of these rules and rules which are believed to be related will be included in the discussion whenever possible. All phonemes are given in IPA symbols; a dash (-) is a place-holder for a letter which is to be converted in a later stage. The result of application of stress rules (to be discussed later) is given without comment following each derivation.

A.	PTOLEMAIC	Input
	PTOLEMA + IC	Result of Stage 1
	t-l-m- + -	Result of Stage 3
	talɛme + ɪk	Result of Stage 3
	$\overset{2}{\text{tal}}\overset{1}{\text{ɪ}}\text{me+ɪk}$	Final result after stress --primary stress appears over the /e/ and secondary stress over the /a/

In the first stage, [ic] is recognized as a suffix and a plus (+) is inserted to its left. Since no other affixes are recognized, Stage 1 is terminated.

Morph-initial [pt] is pronounced /t/, and [l] and [m] are given the pronunciations /l/ and /m/ respectively, according to the most general rule in the rule sequence for each. (The most general rule is the final rule in the rule sequence and contains no specified context.)

In Stage 3, the contexts of the vowels [o] and [e] are not among those contexts specified in the sequence of rules and are pronounced according to the final, context-independent rule. The vowel [a], on the other hand, precedes another vowel and, for this reason is lengthened (tensed). The suffix [+ic] is word-final and receives the pronunciation /ɪk/. In the final result, stress rules have been applied and unstressed non-tense vowels have been reduced.

Generalizations and Related Rules

1.) Morph-initial [pm] and [ps] are given the pronunciations /n/ and /s/ respectively, the [p] remaining silent as in morph-initial [pt].

2.) Vowels in pre-vocalic position are usually lengthened (tensed).

3.) There is only one context in which [ic] is pronounced /ɪs/ rather than /ɪk/, i.e., preceding the variable representing the vowels [i], [e] and [y].

B.	TABLE	Input
	TABLE	Result of Stage 1
	t-b .1	Result of Stage 2
	teb 1	Result of Stage 3

1 teb .1	Final result after stress

The result of Stage 1, in this case, is the same as the input since no affixes are detected.

The letter [t] is pronounced by the most general rule in its rule sequence, and [b] has only one given pronunciation. However, [l] precedes morph-final [e] and is itself preceded by another consonant, [b]. In this context, [l] is syllabic.

The sequence [ble] now forms a very specific context for the third stage. . The letter [a] when followed by [Cle] is lengthened if the consonant, C, is neither [r] nor [l]. The vowel [e] is morph-final and therefore silent.

Generalizations and Related Rules

1.) The rules for [bt] and [mb] in which the [b] is silent are sequenced preceding the single rule for [b].

2.) All vowels except [e], if located in the first syllable of a morph, are long if followed by [Cle#] where C is neither [r] nor [l]. Examples are "maple," "bible," "ogle" and "bugle." An exception is "triple." The letter [e] apears to be long in this context only if it is part of a vowel digraph, e.g., the vowel in "treble" is short, but the vowel digraphs in "eagle," "people" and "beetle" are long. Vowels in this context which do not appear in the first syllable must be converted to short pronunciations so that they will not be given primary stress by the stress rules, e.g., "monocle," "barnacle."

C.	CARIBOU	Input
	CARIBOU	Result of Stage 1
	k-r-b--	Result of Stage 2
	kəɚrɪbu	Result of Stage 3
	1	
	kəɚrību	Final result after stress

During Stage 1, no affixes are detected. Converting consonants in Stage 2, we find that [r] is pronounced according to the most general rule in its rule sequence and that [b] has only one given pronunciation. The letter [c], because it precedes [a], is pronounced /k/.

When [a] precedes [r] which, in turn, precedes either a vowel or another [r] within the same morph, it usually has the pronunciation /əɚ/. The letter [i], following its most general pronunciation, is assigned the phoneme /ɪ/. Morph-final [ou] is given the pronunciation /u/.

Generalizations and Related Rules

1.) The letter [r] is syllabic if preceded by a consonant other than [r] and followed by a morph-final [e], (e.g., "acre"), or the inflectional suffixes [+s] or [+ed].

2.) The letter [c] is palatalized in some cases, e.g., "special," (context: [V__iV]) "ancient," (context: [n__iV]). It is assigned the phoneme /ʃ/ later in its rule sequence if it is followed by [e], [i] or [y]. It may be noted that this is the same context which assigns the pronunciation /ɪs/ to the suffix [+ic]. If [c] is followed by [a], [o] or [u], it is usually pronounced /k/, as in this example..

3.) When [a] precedes [r], and [r] is not followed by either a vowel or another [r] within the same morph, [a] is pronounced /a/, (e.g., "far," "cartoon") unless preceded by the phoneme /w/, e.g., "warble," "warp," "war," "wharf," "quarter").

4.) In a word such as "macaroon," the [a] preceding [rV] is assigned pronunciation /ɔɻ/ in the phonological rules and is reduced to schwa in the stress rules because it is unstressed.

D.	SCENARIO	Input
	SCENARIO	Result of Stage 1
	ss-n-r--	Result of Stage 2
	ssɛnɔɻrio	Result of Stage 3
	----------	----------
	sɪnɔɻrio	Final result after stress

In Stage 2 we find that the consonant cluster [sc], like the letter [c], usually has the sound of /s/ preceding [e], [i] or [y]. The letter [r] does not occur in any context given in its rule sequence and is therefore given its most general pronunciation. There is only one rule for the pronunciation of [n].

Moving on to Stage 3, the vowel [e] receives the pronunciation /ɛ/ given by its most general rule. The vowel [a] follows the rule given in the previous example. The vowel [o] is morph-final and has the feature [-constricted pharynx], and is lengthened accordingly. Because the vowel [i] precedes another vowel, it is lengthened also.[13]

Generalizations and Related Rules

1.) The consonant cluster [sc] is given the representation of a

double phoneme because the information that it is orthographically a double consonant is needed both in the vowel rules and in the rules for stress; it is later reduced to a single phoneme.

2.) Two other contexts or [sc] must be ordered before the rule which applied to "scenario." If [sc] precedes [i] followed by another vowel, and certain letters precede [sc], a palatalization effect is observed. When preceded by a vowel in this context, [sc] becomes /ʃ/, e.g., "prescience"; when preceded by an [n], it becomes /č/ or /tʃ/, e.g., "conscious."

3.) The pronunciation [sc] receives in "scenario" is also found preceding syllabic [l] in example B.

4.) If none of the contexts mentioned in 2.) or 3.) are found, the phonemic representation of [sc] becomes /sk/.

5.) The reduction of /ɛ/ to /ɪ̄/ occurs in the stress rules.

E. SUBVERSION — Input

SUB = VERS + ION — Result of Stage 1

---- = v-rž + --- — Result of Stage 2

səb = vʌrž + ən — Result of Stage 3

---

səb=v $\overset{2}{ʒ}\overset{1}{ž}$+ən — Final result after stress

In this example, the suffix [+ion] and the prefix [sub=] are recognized in Stage 1.

There is only one pronunciation provided for the consonant [v], and [r], because it does not fit a specified context for syllabic [r] is given the standard pronunciation. The letter [s] is followed by the sequence [+iV],

making it a candidate for palatalization. The palatalization rule which applies assigns the phoneme /ž/

In the final stage of letter-to-phoneme conversion, the affixes and vowels are considered. The prefix [sub=] has only one possible pronunciation. The letter [e], because it precedes the sequence [rC] where the consonant, C, is not an [r], is given the pronunciation /ʌ/. The palatal phoneme /ž/ now forms a left context for the suffix [+ion], which, being word-final, is pronounced /ən/.

Generalizations and Related Rules

1.) Because [+s] is marked as occurring in word-final position only, the [s] preceding [+ion] is not recognized as a suffix. This step also prevents the [er] preceding the [s] from consideration as a possible suffix.

2.) When an [s] preceding the sequence [+iV] or [iV] is preceded by either a vowel or an [r], it is usually pronounced /ž/. Some examples are "revision," "artesion," "Persian" and "dispersion"; two exceptions are "controversial" and "torsion." When [s] is preceded by [l], and when it occurs as part of the consonant cluster [ss], the phoneme preceding the vowel sequence is /ʃ/, e.g., "emulsion," "Russian." A third pronunciation is observed when [s] is preceded by [n], e.g., "transient," "comprehension."

3.) The sequence /ʌr/ is later changed to /ɜ/.

4.) The sequence [ion] following a non-palatalized consonant is pronounced /iən/, e.g., "oblivion," "criterion," "champion."

5.) The suffix [+ion] may be given other pronunciations if not morph-final. For example, it is pronounced /+iən/ in ganglionic" and "histrionic."

F.	SCIENCE	Input
	SCI + ENCE	Result of Stage 1
	ss- + ----	Result Stage 2
	ssə + ɪns	Result of Stage 3
	s əˈ + ɪns	Final result after stress

In Stage 1, [+ence] is recognized as a suffix. The consonant cluster [sc] precedes [i] and is therefore given the pronunciation /ss/, later changed to /s/ as described earlier. The pronunciation of [i] preceding a vowel as /ay/ is a consequence of its left context being a morph-initial consonant cluster.

Generalizations and Related Rules

1.) Although [+ience] is a possible suffix, it is not recognized as such in this case because of the requirement that at least one consonant and one vowel remain in the "root." This stipulation forces the correct suffix, [+ence], to be recognized.

2.) Because [sc] is morph-initial, it is not palatalized even though it precedes the sequence [iV].

3,) The letter [i] is also pronounced /ay/ if it is followed by a vowel and is morph-initial, e.g., "iotə," "ɪambic."

NEUTRALIZATIONS

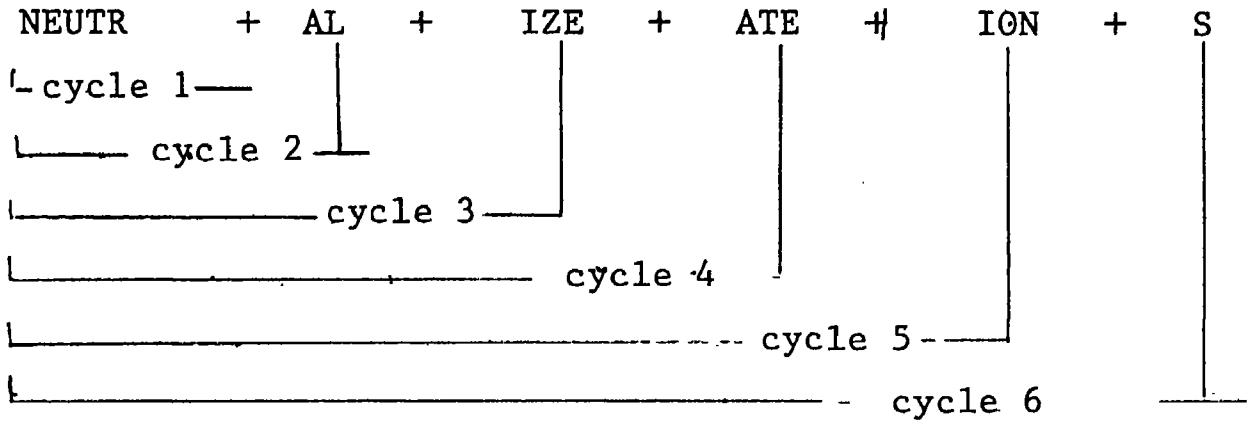


Figure 2

Cyclic Rules (First Phase)

Domain of Application

NOTATION

V	vowel
C	consonant
C_i	at least i consonants
C^j	at most j consonants
C_i^j	at least i consonants; at most j consonants
X,Y,	variables
W	a weak syllable, i.e. a short vowel followed by no more than one consonant (a syllable begins with a vowel and terminates (a) immediately before the next vowel or (b) immediately before a formative boundary if one occurs before the next vowel
[ ]	a feature, e.g., [-long], [1 stress], or a phoneme with specified feature(s), e.g., $\begin{bmatrix} V \\ 1 \text{ stress} \end{bmatrix}$.
$\begin{Bmatrix} A \\ B \\ \vdots \\ P \end{Bmatrix}$	either A or B or ... or P
( )	optional element; material in parentheses is neglected if and only if it does not correspond to context in the word under consideration -- word context is compared with rule context by first comparing it with the maximum string in the rule, i.e., with all parentheses removed, and then by ignoring parenthesized material beginning at the innermost parentheses and proceeding to the outermost parentheses
⟦ ⟧	domain of rule -- formative boundaries of string under consideration for cyclic rules, word-boundaries for last cycle and for non-cyclic rules
$(\ )_a\ (\ )_b$	subscripts making appearance of optional elements conditional (actual condition given below rule)

Figure 3.

$X \rightarrow [y] \; / Y\_\_Z$	assign feature(s) [y] to element X in the context YXZ
$X \rightarrow [y] \; / Y \left[\begin{smallmatrix} \cdot \\ z \end{smallmatrix}\right] Z$	assign feature(s) [y] to an element X with specified feature(s) [z] in the context YXZ

Figure 3 (concluded)

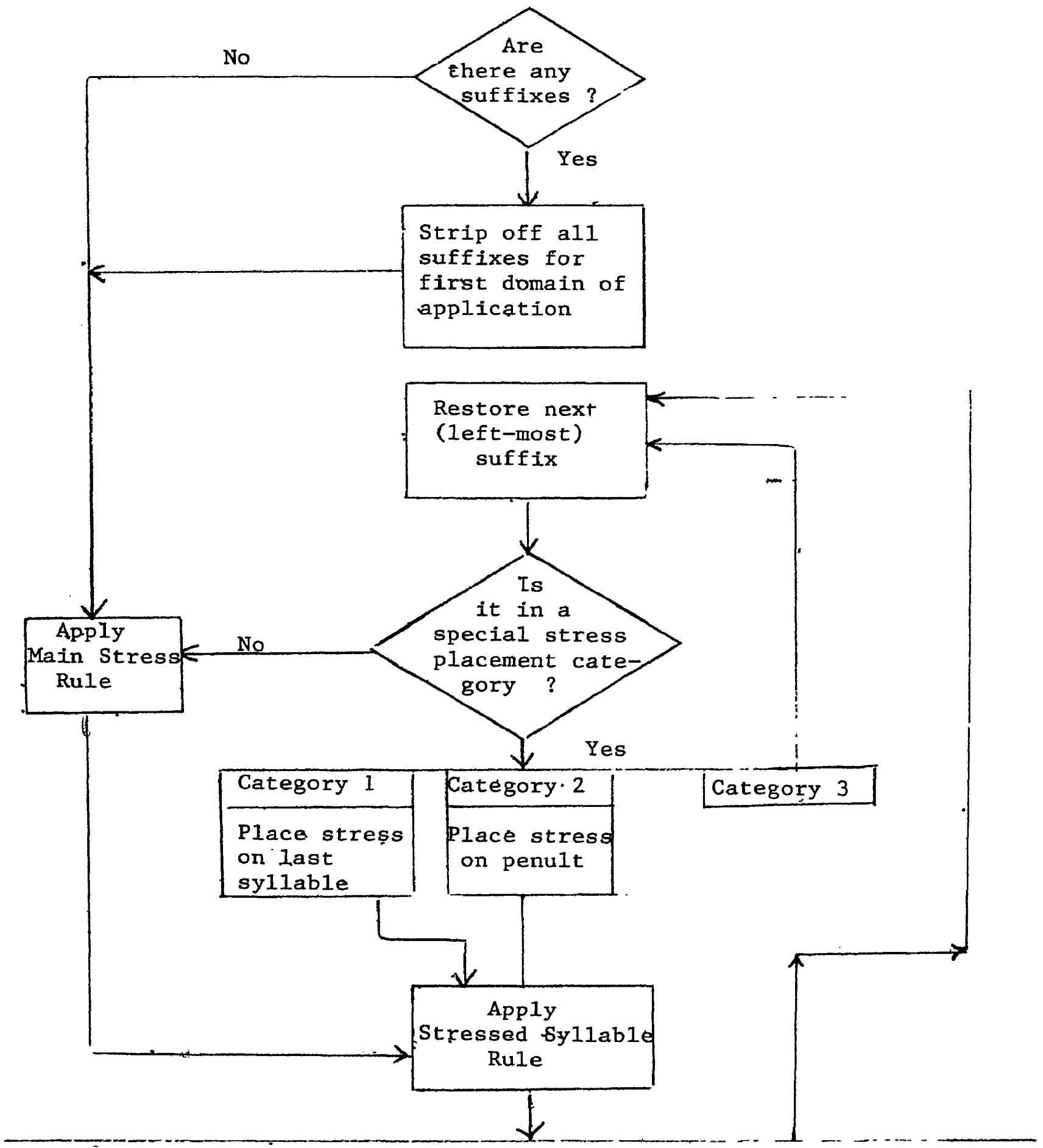


Figure 4.

Stress Rules - Flow Chart

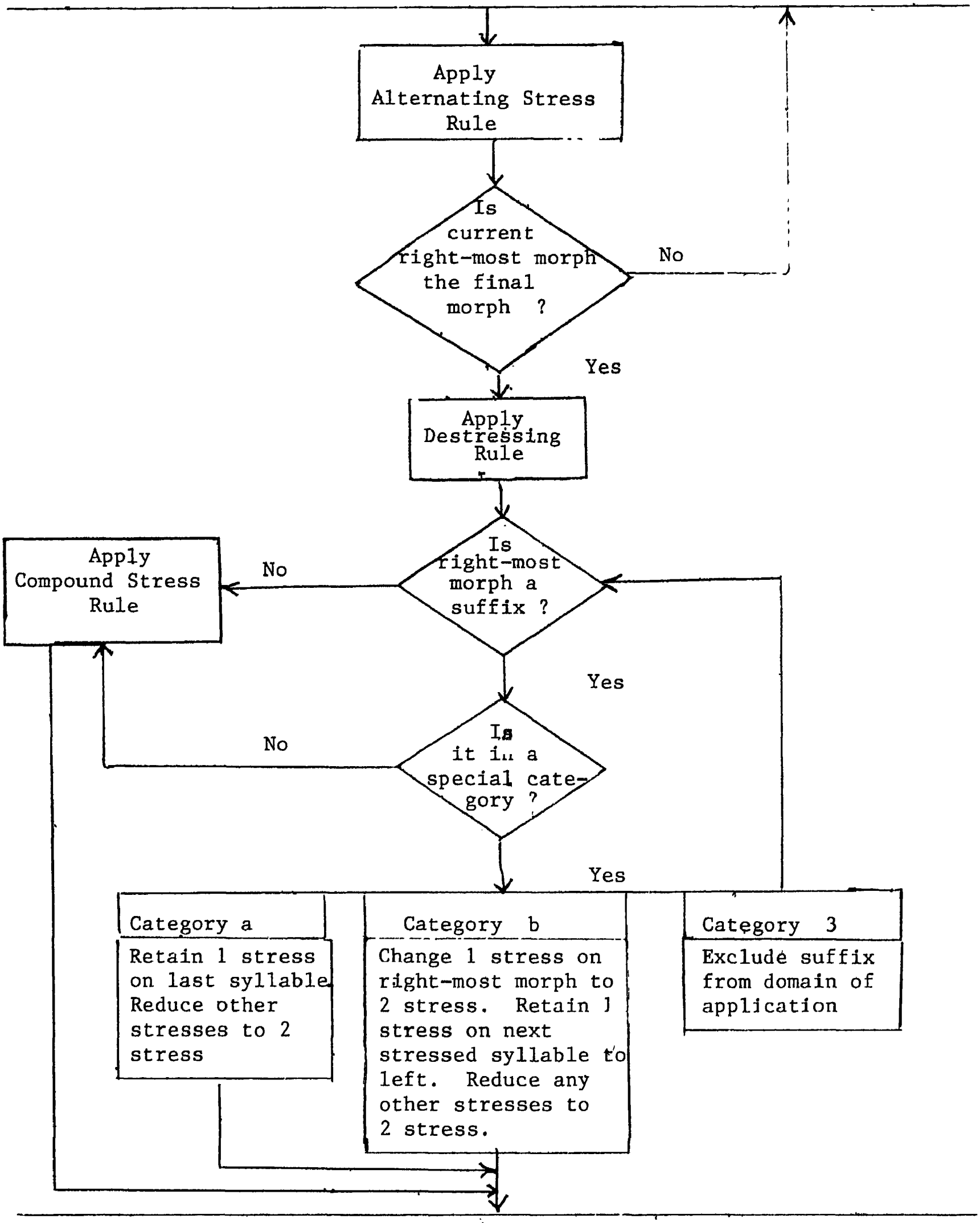


Figure 4 (continued)

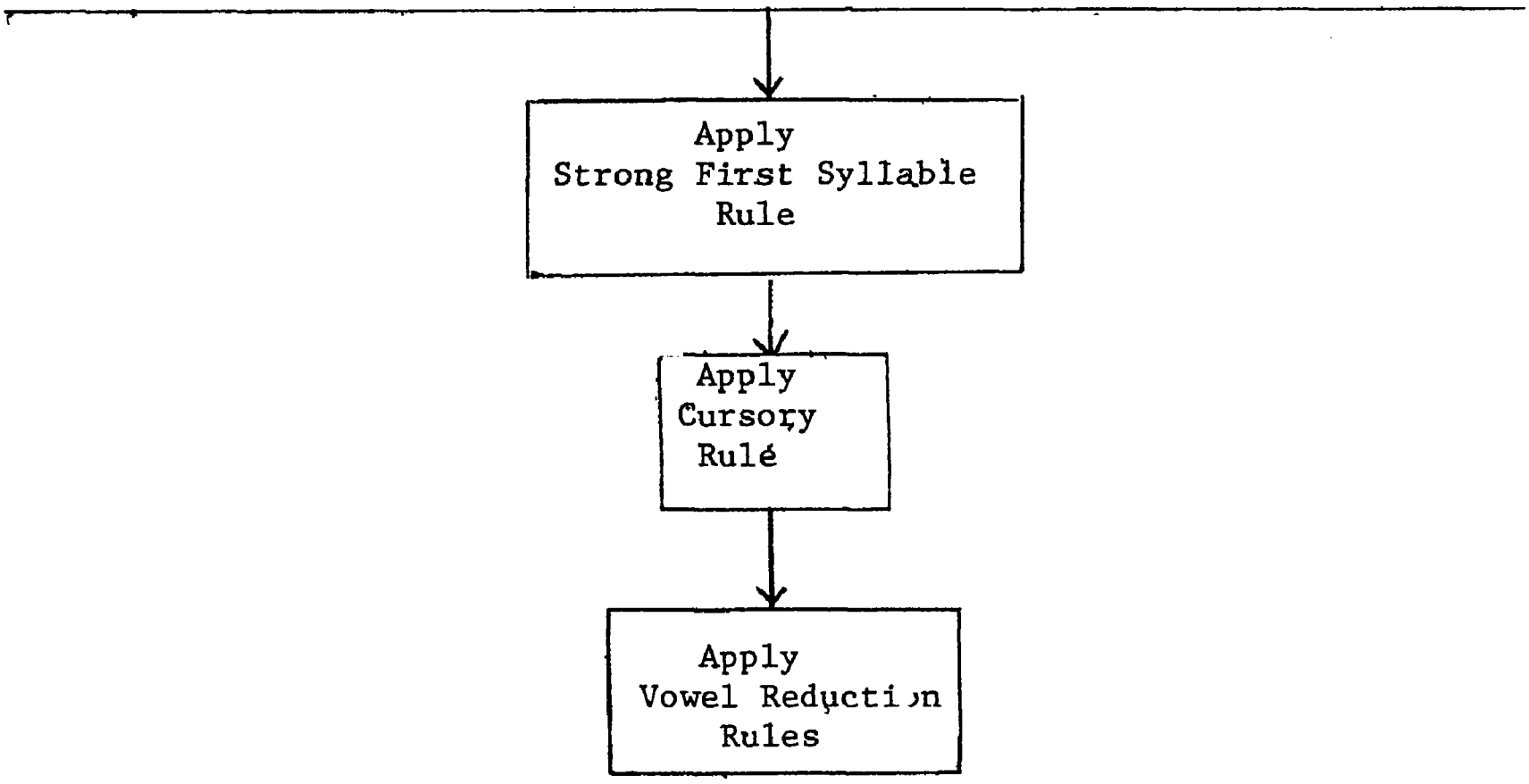


Figure 4. (concluded)

Main Stress Rule (cyclic)

$$V \rightarrow [1\ \text{stress}] \,/\, [\![X\_\_C_0 \left(\left(\begin{Bmatrix} W \\ V \end{Bmatrix} \right) \left\{ \begin{bmatrix} V \\ -\text{long} \\ V \end{bmatrix} C_0 \right\})]\!]$$

Conditions: (1) no stress placement to the left of a prefix boundary
(2) if right-most morph is a suffix, test for special stress placement category; assign [1 stress] or skip cycle according to category.

Stressed Syllable Rule (cyclic)

$$V \rightarrow [1\ \text{stress}] \,/\, [\![X\_\_C_0 \left(\left(\left(\begin{Bmatrix} W \\ V \end{Bmatrix} \right) VC_0 \right) \begin{bmatrix} V \\ 1\ \text{stress} \end{bmatrix} Y]\!]$$

Conditions: (1) Y contains no primary stress
(2) no stress placement to the left of a prefix boundary

Alternating Stress Rule (cyclic)

$$V \rightarrow [1\ \text{stress}] \,/\, [\![X\ \ C_0 \ (V)\ VC_0 \begin{bmatrix} V \\ \overline{1\ \text{stress}} \end{bmatrix} \ C_0]\!]$$

Destressing Rule (non-cyclic)

$$V \rightarrow \begin{bmatrix} -\text{long} \\ -\text{stress} \end{bmatrix} \,/\, [\![C_0 \ (VC_0X)_a \ \begin{bmatrix} \overline{\quad} \\ (-\text{long})_b \end{bmatrix} C \begin{bmatrix} V \\ +\text{stress} \end{bmatrix} Y]\!]$$

Conditions: (1) if $(\ )_a$ is not present, $(\ )_b$ must be present
(2) not applied to the first vowel if applied to second vowel

Compound Stress Rule (non-cyclic)

$$\begin{bmatrix} V \\ 1\ \text{stress} \end{bmatrix} \rightarrow [1\ \text{stress}] \,/\, [\![X\_\_Y \ (VC_0 \left(\begin{bmatrix} /i/ \\ -\text{stress} \end{bmatrix} \right))]\!]$$

Conditions: (1) Y contains no 1 stress
(2) if right-most morph is a suffix, check for special stress retention or stress exclusion category and reassign [1 stress] according to category.

Figure 5.

Stress Placement Rules

Strong First Syllable Rule (non-cyclic)

$$V \rightarrow [2\ \text{stress}] \,/\, [\![C_0 \left[\frac{\quad}{(+\text{long})_a} \right] (C_2)_b \ Y]\!]$$

Condition: (1) $(\quad)_a$ or $(\quad)_b$ must be present

Cursory Rule (non-cyclic)

$$V \rightarrow \begin{bmatrix} -\text{long} \\ -\text{stress} \end{bmatrix} \quad / [\![X \begin{bmatrix} V \\ 1\text{stress} \end{bmatrix} C_0 \_\_ CV \ Y]\!]$$

Condition: (1) if right-most morph is a suffix, check for stress exclusion category

Vowel Reduction Rules (non-cyclic)

$$\begin{bmatrix} \{/\text{ɪ}/ \\ /\varepsilon/\} \\ -\text{stress} \end{bmatrix} \rightarrow \text{ɪ} / \ X \_\_ Y$$

$$\begin{bmatrix} V \\ -\text{long} \\ -\text{stress} \end{bmatrix} \rightarrow \partial / \ X \_\_ Y$$

Figure 5. (concluded)

III. Lexical Stress Placement

The stress rules which have been implemented are a modification of a set of ordered rules developed by Halle.[14] Modifications fall into three categories: (1) adjustments due to the condition that input is completely phonemic, (2) reduction of the number of stress levels to 1 stress (primary) 2 stress (stress less than primary) and 0 stress, and (3) addition of special suffix-dependent stress categories.

Application of the rules proceeds in two phases. The first phase consists of the application of three ordered rules which are applied cyclically, first to the root, then to the root and left-most suffix combined. The process continues with one more suffix adjoined to the string under consideration before each cycle begins until the end of the word is reached. This cyclic phase is devoted solely to the placement of primary stress. The second, non-cyclic phase, includes the application to the entire word of ordered rules and reduces all but one of the primary stress marks to secondary or zero stress.

In the following section, stress placement rules will be given in symbolic form. Each rule which contains more than one case is broken down into cases for which brief descriptions and examples are given. The rules are listed in the order in which they apply and are marked either "cyclic" or "non-cyclic." Particular modifications to each rule will be given at the end of the discussion about that rule under the subheading Modifications. (See Figure 3 for an explanation of notation, Figure 4 for a flow chart of the stress rules and Figure 5 for the complete set of stress-placement rules in linguistic notation.

Main Stress Rule (cyclic)

$$V \rightarrow [1\ \text{stress}] \ / [[X\_\_C_0 \ (\left(\begin{Bmatrix} W \\ V \end{Bmatrix}\right)\left\{\begin{bmatrix} V \\ -\text{long} \\ V \end{bmatrix} C_0\right\})]]$$

Condition: (1) no stress placement to the left of a prefix boundary
(2) if right-most morph is a suffix, test for special stress placement category ; assign [1 stress] or skip cycle according to category.

Case 1. (Maximum string; all parentheses removed)

$$V \rightarrow [1\ \text{stress}] \ / [[X\_\_C_0 \begin{Bmatrix} W \\ V \end{Bmatrix}\left\{\begin{bmatrix} V \\ -\text{long} \\ V \end{bmatrix} C_0\right\}]]$$

(a) Assign 1 stress to the vowel in a syllable preceding a weak cluster followed by a morph-final syllable containing a short vowel and zero or more consonants:

difficult	dɪ1ffɪkʌlt

(b) Assign 1 stress to the vowel in a syllable preceding a weak cluster followed by a morph-final vowel:

oregano	ɔrɛ1gɚno

(c) Assign 1 stress to the vowel in a syllable preceding a vowel followed by a morph-final syllable containing a short vowel and zero or more consonants:

secretariat	sɛkrɛ1tɚriɚt

(d) Assign 1 stress to the vowel in a syllable preceding a vowel followed by a morph-final vowel:

oratorio	ɔrɚ1tɔrio

Case 2. (Innermost parenthesized string excluded)

$$V \rightarrow [1\ \text{stress}] \ / \ [[X\_\_C_0 \left\{ \begin{bmatrix} V \\ -\text{long} \\ V \end{bmatrix} C_0 \right\}]]$$

(a) Assign 1 stress to the vowel in a syllable preceding a short vowel and zero or more consonants:

edit	ɛdɪt
bitumen	baɪtumɛn

(b) Assign 1 stress to the vowel in a svllable preceding a morph-final vowel:

agenda	əjɛndə

Case 3. (All parenthesized strings excluded)

$$V \rightarrow [1\ \text{stress}] \ / \ [[X\_\_C_0]]$$

(a) Assign 1 stress to the vowel in the last syllable:

stand	stənd	go	go
parole	pərol	hurricane	hʌrɪken

(reduced to 2 stress by a later rule)

Conversion of the Main Stress Rule into algorithmic form is facilitated by ordering the above cases in the following manner:

Algorithmic Order of Application: (1) If the final syllable is the only syllable, or if it consists of a long vowel followed by at least one consonant, the final vowel receives primary stress. Otherwise, (2) if there are only two syllables, or if the penultimate syllable terminates in more than one consonant or if it consists of a long vowel followed by at least one consonant, the penultimate vowel receives primary stress. Otherwise, (3) the antepenultimate vowel receives primary stress.

Modifications: The presence of the optional vowel immediately preceding another vowel and the presence of the morph-final vowel are necessary

modifications of the Main Stress Rule due to the difficulty of retrieving the long (tense) pronunciation of a laxed vowel when its orthographic representation is no longer available.

The Main Stress Rule, as developed by Halle, applies only to roots which function as nouns and to suffixed forms. However, until parsing methods are further developed, it will not be possible to take advantage of known parts of speech.[15] For this reason, the Main Stress Rule is currently applied to all roots.

The suffixes referred to in Condition (2) fall into two categories. Some suffixes are marked to force stress to be placed on either the final or the penultimate syllable of the root and suffixes under consideration. This placement of stress replaces the MSR on the cycle in which the special suffix is the right-most morph. These suffixes are listed below with the phonemic representation which actually appears as input.

	Example
EE - /ii/, final-syllable stress, retained by special categorization	2 1 trainee
EER - /ɪr/, final-syllable stress, retained by special categorization	2 1 buccaneer
ESCE - /ɛ̄s/, final-syllable stress, retained by special categorization	2 1 luminesce
ESQUE - /ɛsk/, final-syllable stress, retained by special categorization	2 1 arabesque
ETTE - /ɛt/, final-syllable stress, retained by special categorization	2 1 marionette
OON - /un/, final-syllable stress, retained by special categorization	1 spitoon

SELF - /sɛlf/, final-syllable stress, retained by special categorization	2 1 herself
FUL - /fʊl/, final-syllable stress, later reduced by Compound Stress Rule	1 2 bushelful
HOOD - /hʊd/, final-syllable stress, later reduced by Compound Stress Rule	1 2 womanhood
IFY - /ɪfəɻ/, final-syllable stress, later reduced by Compound Stress Rule	2 1 2 humidify
IZE - /əɻz/, final-syllable stress, later reduced by Compound Stress Rule	1 2 radicalize
OID - /ɔɻd/, final-syllable stress, later reduced by Compound Stress Rule	1 2 ovoid
SHIP - /ʃɪp/, final-syllable stress, later reduced by Compound Stress Rule	1 2 friendship
ISM - /ɪzəm/, penultimate-syllable stress, later reduced by special categorization	2 1 2 romanticism
ARY - /əɹi/, /ɛri/, /ɛrɪ/, penultimate-syllable stress reduced by Compound Stress Rule	2 1 circumlocu- 2 tionary
penultimate-syllable stress, deleted by Cursory Rule	2 1 infirmary
ORY - /ɔri/, /ɔrɪ/, penultimate-syllable stress, reduced by Compound Stress Rule	2 1 2 inhibitory
penultimate-syllable stress, deleted by Cursory Rule	1 refractory
ERY - /ɛri/, penultimate-syllable stress, reduced by Compound Stress Rule	1 2 stationery
penultimate-syllable stress, deleted by Cursory Rule	1 slippery
ATORY - /ətɔri/, penultimate-syllable stress, reduced by Compound Stress Rule	1 2 systematory
ITION - /ɪʃən/, penultimate-syllable stress, retained	1 sedition

IFIC - /ɪfɪk/, /ɪfɪs/, penultimate-syllable stress, retained — 1 specific

penultimate-syllable stress, reduced by Destressing Rule — 2 1 specificity

IC - /ɪk/, /ɪs/, penultimate-syllable stress, retained — 2 1 orthographic

penultimate-syllable stress, reduced by Compound Stress Rule — 2 1 simplicity

The other category of suffixes referred to in Condition (2) does not affect stress; the cycle in which such a suffix is right-most in the domain is skipped. Later cycles, however, do include the suffix as part of their domain of application. These suffixes are listed below, and are accompanied by examples demonstrating their inclusion in this category.

ABLE: (a) all words terminating in ICABLE, e.g., 1 eradicable, 1 communicable

(b) 1 formidable, 1 noticeable, 1 manageable, 1 knowledgeable

ABLY: as with ABLE above

AGE: (a 1 brigandage, 1 2 vagabondage, 1 2 chaperonage

(b) 1 2 anecdotage -- at the time Walker's rhyming dictionary was compiled, the two stresses were interchanged

DOM: (a) 1 bachelordom

(b) words such as 1 christendom and 1 martyrdom do not supply evidence; "dom" must be considered a separate syllable, i.e., the syllable preceding "om" is not strong

ED: (a) 1 2 opinionated, 1 talented, 1 shepherded

(b) Exceptions occur in words with no secondary stress and with primary stress more than two syllables to the left before the affixation of +ed, e.g.,
1 2 1 2
precedented, interested (in some dialects)

(c) Note that ED has no vowel in its pronunciation if not preceded by [t] or [d], i.e., it is not a separate syllable.

EER: (a) caravaneer^{2}, charioteer^{1} ... charioteer^{2}

EN: There is no evidence of stress change due to the suffixation of "en"; most words to which it is added are 1-syllable roots.

ER: countenancer^{1}, circumscriber^{1} ... circumscriber^{2}

ESQUE: Raphaelesque^{2}, harlequinesque^{1} ... harlequinesque^{2}

ES: (a) privileges^{1}, cartilages^{1}, luridnesses^{1}

(b) impoverishes^{1}

(c) Note that ES (and S affixed to morph-final [e]) has no vowel in its pronunciation if not preceded by s,z,š,č,ž, or ǰ.

EST: There is no evidence of stress change due to the suffixation of "est."

ETH: seventieth^{1}

FUL: No evidence of stress change

HOOD: parenthood^{1} ... parenthood^{2}

IBLE: (a) eligible^{1}, intelligible^{1} ... intelligible^{2}

(b) words such as putrescible^{1} and fermentescible^{1} are not exceptions; the verbal ending ESCE always carries primary stress

IBLY: as with IBLE above

ILE: replicatile^{1}, fluviatile^{1} ... fluviatile^{2}

ING: (a) conveyancing^{1}, countenancing^{1} ... countenancing^{2}

(b) Exceptions may occur in those contexts mentioned under ED. In the case of "countenancing," the syllable consisting of "en" is generally so reduced that it is imperceptible as a syllable.

ISH: (adjectival) amateurish^{1}, sycophantish^{1} ... sycophantish^{2}

1 2 1 2
ISM: (a) Pharisaism, Sadduceeism
1 2 21 2
(b) invalidism, theatricalism
1 2 1 2
(c) vagabondism, monarchism
1 2 1 2 1 2
IZE: (a) standardize, jeopardize, energize
1 2 1 2 1 2 1 2
(b) radicalize, memorialize, secularize, proselytize
1 1 1
LESS: conscienceless, characterless, objectless

LET: No evidence of stress change due to the affixation of "let."
1 1 1
LY: particularly, passionlessly, precipitously
1 1

MENT: (a) Words such as government and sojournment indicate that MENT should be placed in this category.

(b) Most words of four or more syllables are given alternate pronunciations corresponding to the placement of MENT in either this category or in the category of regular stress placement, e.g.,
1 1 1 1
advertisement / advertisement, medicament / medicament
1 1
NESS: disinterestedness, haphazardness
1 1 1 2
OR: governor, warrantor, incubator
1 1 1
RY: heraldry, wizardry, charlantanry

SELF: No evidence of stress change
1 2
SHIP: (a) umpireship
1 2 1 2 1 2 1 2
(b) advocateship, candidateship, laureateship, cardinalship

SOME: No evidence of stress change
1 1
TY: sheriffalty, suzerainty
1 1 1
URE: judicature, triplicature, caricature

All other suffixes not in the above categories receive stress according to the general form of the Main Stress Rule.

Stressed Syllable Rule (cyclic)

$$V \rightarrow [1\ \text{stress}] \quad / [[X\_\_C_0 \quad ((\left(\begin{Bmatrix} W \\ V \end{Bmatrix}\right) VC_0) \begin{bmatrix} V \\ 1\ \text{stress} \end{bmatrix}) \quad Y]]$$

Conditions: (1) Y contains no primary stress
(2) no stress placement to the left of a prefix boundary

Case 1. (Maximum string: all parentheses removed)

$$V \rightarrow [1\ \text{stress}] \quad / \quad [[X\_\_C_0 \begin{Bmatrix} W \\ V \end{Bmatrix} VC_0 \begin{bmatrix} V \\ 1\ \text{stress} \end{bmatrix} \quad Y]]$$

(a) Assign 1 stress to the vowel in a syllable preceding a weak cluster followed by a vowel and any number of consonants which is followed by the right-most primary-stressed vowel:

oxygenate — ① 1
aksɪjɛnet
(stress on final syllable later reduced)

(b) Assign 1 stress to the vowel in a syllable preceding a vowel followed by a vowel and any number of consonants which is followed by the right-most primary-stressed vowel:

stereobate — ① 1
stɛriobet
(stress on final syllable later reduced)

Case 2. (Innermost parenthesized string excluded)

$$V \rightarrow [1\ \text{stress}] \quad / \quad [[X\_\_C_0\ VC_0 \begin{bmatrix} V \\ 1\ \text{stress} \end{bmatrix} \quad Y]]$$

(a) Assign 1 stress to the vowel two syllables to the left of the right-most primary-stressed vowel:

propaganda — 1 1
prapəgændə
(stress on left-most stressed vowel later reduced)

Case 3. (Next innermost parenthesized string excluded)

$$V \rightarrow [1\ \text{stress}] \quad / \quad [[X\_\_C_0 \begin{bmatrix} V \\ 1\ \text{stress} \end{bmatrix} \quad Y]]$$

(a) Assign 1 stress to the vowel one syllable to the left of the right-most primary-stressed vowel, i.e., to the vowel in the first syllable of the root:

Hormone — ① 1
hɔrmon

Case 4. (All parenthesized strings excluded)

$$V \rightarrow [\ 1\ \text{stress}\ ] \ / \ [\ \ X \_\_ C_0\ ]$$

(a) Assign 1 stress to the vowel in the last syllable, i.e., to the vowel in the only syllable of the root:

stand → st$\acute{a}$nd

(assigning 1 stress to a vowel which already carries 1 stress has no effect unless the rule specifies as in the Compound Stress Rule, that the vowel must previously be 1-stressed.)

Algorithmic Order of Application: (1) If the right-most syllable containing primary stress is the left-most syllable in the word, no stress is assigned. Otherwise, (2) if the syllable preceding the right-most stressed syllable is the only syllable preceding it, assign primary stress to the vowel in that syllable. Otherwise, (3) if the second syllable to the left of the right-most stressed syllable is the left-most syllable, or if it terminates in more than one consonant or consists of a long vowel followed by at least one consonant, assign primary stress to the vowel in that syllable. Otherwise, (4) the vowel in the third syllable to the left of the right-most stressed syllable receives stress.

Modifications: The optional vowel in pre-vocalic position appears in the Stressed Syllable Rule as well as in the Main Stress Rule. Its presence prevents words such as "stereobate," "alveolate," and "heliotrope" from being stressed incorrectly.

The Stressed Syllable Rule, as developed by Halle, places stress on the final syllable of the non-nouns which have been excluded from the domain of application of the Main Stress Rule. Words for which the categorization of noun/non-noun appear to be most useful are those in which a one-syllable

prefix precedes a one-syllable root or bound morpheme, e.g., $[^{1\ 2}_{N}$permit$]_N$ vs. $[^{1}_{V}$permit$]_V$, $[^{1\ 2}_{N}$insult$]_N$ vs. $[^{1}_{V}$insult$]_V$ Because there are many more verbs of this sort than nouns, the Stressed Syllable Rule has been modified to prevent the retraction of stress into a prefix. The effect of this modification is to produce only the verbal pronunciation of two-syllable noun/verb pairs. Another more positive, effect is the correct placement of stress in verbs such as ^{1}edit, ^{1}inhibit and ^{1}pummel. However, two-syllable nouns of the form "prefix-root" which have no verbal counterpart are stressed incorrectly, e.g. ^{1}empire, ^{1}inverse. (This modification will be removed or changed after a parsing algorithm is incorporated in the system.)

Alternating Stress Rule (cyclic)

$$V \rightarrow [1\ \text{stress}] \ / [[X\_\_\_C_0\ (V)\ VC_0\ [^{V}_{1\ \text{stress}}]\ C_0]]$$

Case 1. (Maximum string)

$$V \rightarrow [1\ \text{stress}] \ / [[X\_\_\_C_0VVC_0\ [^{V}_{1\ \text{stress}}]\ C_0]]$$

(a) Assign 1 stress to the vowel three syllables to the left of a primary-stressed vowel occurring in the last syllable if the following syllable contains only a vowel:

heliotrope — ①^{1}hiliotrop

(stress in last syllable later reduced)

Case 2 (Parenthesized string excluded)

$$V \rightarrow [1\ \text{stress}] \ / [[X\_\_\_C_0VC_0\ [^{V}_{1\ \text{stress}}]\ C_0]]$$

(a) Assign 1 stress to the vowel two syllables to the left of a primary-stressed vowel occurring in the last syllable:

gelinate — ^{1}ʤɛl^{1}ɪ^{1}neɪt

(stress in first syllable later deleted; stress in last syllable later reduced)

Algorithmic Order of Application: (1) If there are at least two syllables preceding a primary-stressed vowel in the last syllable of the phoneme string, and if the first of these two syllables is composed of more than a single vowel, place primary stress on the vowel two syllables to the left of the vowel with primary stress. Otherwise, (2) if there are at least three syllables to the left, the second of which is composed of a single vowel, place primary stress on the vowel three syllables to the left of the vowel with primary stress. Otherwise, (3) no stress assignment is made.

Exceptions: Note that words such as peregrinate, oxygenate and metropolitanate which are correctly stressed by the Stressed Syllable Rule are stressed incorrectly, thereafter, by the Alternating Stress Rule.

Modification: The optional vowel in pre-vocalic position appears in the Alternating Stress Rule as well as in the Main Stress and Stressed Syllable rules.

Proposed Modification: The restriction of the Alternating Stress Rule to words in which a prefix boundary does not precede the final primary-stressed syllable could be constrained to verbs. Such a constraint would provide the correct stress placement in nouns and adjectives such as multiform, contraband, intercept and miniskirt while retaining correct stress placement in the verbs intercept, contradict and comprehend. Such a modification would require moving the Strong First Syllable Rule in Halle's scheme to follow the Compound Stress Rule, assigning [2 stress] in the same context in which [1 stress] was previously assigned. This modification has already been implemented in this program for independent reasons.

and is discussed under the heading Modifications in the Strong First Syllable Rule.

Destressing Rule (non-cyclic, applicable to all vowels having required context)

$$V \rightarrow \begin{bmatrix} -\text{long} \\ -\text{stress} \end{bmatrix} \; / \; [\; C_0 \; (VC_0X)_a \; \left[(\overline{-\text{long}})_b\right] \; C \; \begin{bmatrix} V \\ +\text{stress} \end{bmatrix} \; Y \;]$$

Conditions: (1) if $(\quad)_a$ is not present, $(\quad)_b$ must be present
(2) not applied to first vowel if applied to second vowel

Case 1. (Vowel to be reduced not in first syllable)

$$V \rightarrow \begin{bmatrix} -\text{long} \\ -\text{stress} \end{bmatrix} \; / \; [\; C_0VC_0X\_\_C \; \begin{bmatrix} V \\ +\text{stress} \end{bmatrix} \; Y \;]$$

(a) Shorten and destress any vowel not in the first syllable which is followed by a single consonant and a stressed vowel:

instrumental — ɪnstrʊmɛntəl
(/u/ reduced to /ʊ/, later to /ə/)

Case 2. (Vowel to be reduced is in first syllable)

$$V \rightarrow [-\text{stress}] \; / \; [\; C_0 \; \left[\frac{\quad}{-\text{long}}\right] \; C \; \begin{bmatrix} V \\ +\text{stress} \end{bmatrix} \; Y \;]$$

(a) Destress a non-long vowel in the first syllable which is followed by a single consonant and a stressed vowel:

gelatinate — ǰɛlətɪnet

Algorithmic Application: (1) If a vowel not in the first syllable is immediately followed by one consonant and a vowel which has previously been assigned primary stress, shorten (lax) it if it is long, and remove any stress it has been assigned. (2) If a short vowel is in the first syllable, and is immediately followed by one consonant and a vowel which has previously been assigned primary stress, and if (1) does not apply to the vowel in the second syllable, remove any stress that has been assigned to

the vowel in the first syllable.

Modification: The single required consonant preceding the primary stressed vowel has been changed from C_0^1 (zero or one consonant) ~~to~~ C (~~exactly~~ one consonant) so that pre-vocalic vowels are not shortened.

Compound Stress Rule (non-cyclic)

This rule, as developed by Halle, applies to both compounds and non-compounds. As it applies to words converted by letter-to-phoneme rules in the program, and therefore to non-compounds only, its effect is to locate the primary stress which is to be retained. All other primary stress is reduced to secondary. Halle has used the Nuclear Stress Rule for both phrase-level stress and the reduction of secondary to tertiary stress in lexical items. Neither is necessary in this algorithm; the Nuclear Stress Rule has therefore been omitted.

$$\begin{bmatrix} V \\ 1 \text{ stress} \end{bmatrix} \rightarrow [1 \text{ stress}] \quad / \quad [\![X \_\_ Y \ (VC_0 \ (\begin{bmatrix} /i/ \\ -\text{stress} \end{bmatrix}))]\!]$$

Condition: (1) Y contains no 1 stress
(2) if right most morph is a suffix, check for special stress retention or stress exclusion category and reassign [1 stress] according to category

Case 1. (Maximum string)

$$\begin{bmatrix} V \\ 1 \text{ stress} \end{bmatrix} \rightarrow [1 \text{ stress}] \quad / \quad [\![X \_\_ YVC_0 \ \begin{bmatrix} /i/ \\ -\text{stress} \end{bmatrix}]\!]$$

(a) Retain 1 stress on a vowel if it is followed by at least one syllable and a word-final unstressed /i/. Reduce all other 1 stress to 2 stress:

legendary — lɛjɛndɔɹi

Case 2. (Innermost parenthesized string excluded)

$$\begin{bmatrix} V \\ 1 \text{ stress} \end{bmatrix} \rightarrow [1 \text{ stress}] \quad / \quad [[\, X \_\_\_ YVC_0 \,]]$$

(a) Retain 1 stress on a vowel if it is followed by at least one syllable. Reduce all other 1 stress to 2 stress:

hurricane	hʌr$\overset{1}{\overset{\nearrow 2}{\imath}}$ken	gastritis	g$\overset{\nearrow 2}{\text{æ}}str\overset{1}{\text{ə}}$ tɪs
trinitarian	trɪnɪt$\overset{\nearrow 2}{\text{æ}}\overset{1}{\text{r}}$iən		

Case 3. (All parenthesized strings excluded)

$$[1 \overset{V}{\text{stress}}] \rightarrow [1 \text{ stress}] \quad / [[\, X \_\_\_ Y \,]]$$

(a) Retain 1 stress on the only vowel to which it has been assigned:

stand	st$\overset{1}{\text{æ}}$nd	edit	$\overset{1}{\varepsilon}$dɪt
difficult	$\overset{1}{\text{d}}$ɪffɪkʌlt		

Algorithmic Order of Application: (1) If primary stress occurs only once, no changes are made. Otherwise, (2) if the right-most vowel with primary stress is followed by at least one more syllable, the right-most of which is not composed of an unstressed /i/, it retains primary stress and all other primary stress is reduced to secondary. (3) If the right-most vowel with primary stress is (a) the right-most vowel in the word, or (b) the right-most vowel with the exception of a final syllable composed of an unstressed /i/, the first primary-stressed vowel to its left retains primary stress and all other primary stress is reduced to secondary.

Modifications: As mentioned previously, input to the stress rules from the letter-to-phoneme program does not include compounds. The part of the rule designed for compounds is, therefore, omitted.

This rule formerly contained the letter [y] instead of $\begin{bmatrix} /i/ \\ -\text{stress} \end{bmatrix}$ which has been substituted due to unavailability of the original orthography.

The suffixes referred to in Condition (2) fall into two categories. Those suffixes discussed under Condition (2) of the Main Stress Rule which do not affect stress placement are excepted from the domain of the Compound Stress Rule if they are either word-final or precede another word-final suffix in the same category.

The other category of suffixes is marked for special stress retention. The following suffixes retain primary stress in word-final position under Condition (2) of the Compound Stress Rule:

	2 1 2 1
EE:	trainee, legatee
	2 1 2 1
EER:	buccaneer, engineer
	2 1 2 1
ESCE:	luminesce, acquiesce
	2 1 2 1
ESQUE:	arabesque, Romanesque
	2 1 2 1
ETTE:	marionette, majorette
	2 1 1
OON:	macaroon, baboon

The following suffix does not retain primary stress on the penultimate syllable under Condition (2) of the Compound Stress Rule:

	1 2 2 1 2
ISM:	Babism, Romanticism

Note: This categorization is equivalent to the statement that syllabic M does not function as a syllable in morph-final position. The same stress pattern appears in words ending in [ithm], although it is not included here as a suffix, e.g.,

1 2 1 2
logarithm, algorithm.

The same categorization should be extended to morph-final syllabic [l]. However, it does not function as a suffix, e.g.,

1
corpuscle

The original set of stress rules included the Trisyllabic Shortening Rule at this point in the ordering. The rule was stated as follows:

$$\begin{bmatrix} V \\ 1 \text{ stress} \end{bmatrix} \rightarrow [-\text{long}] \quad / \left[\left[X \_\_ C \begin{bmatrix} \text{'} V \\ -\text{stress} \end{bmatrix} CV \quad Y \right] \right]$$

Condition: (1) does not apply to /u/

Test results indicated mispronunciations arising from its application. A study[16] was undertaken to determine the usefulness of this rule and to uncover problem areas which might lead to a more proper resolution of observed effects for which the Trisyllabic Shortening Rule was formulated. It was found that a restatement of phonological rules, including the requirement of a short vowel in a one-syllable root preceding a single consonant, and certain suffixes, obviated the need for the Trisyllabic Shortening Rule in the set of stress rules.

Strong First Syllable Rule (non-cyclic)

$$V \rightarrow [2 \text{ stress}] \quad / \left[C_0 \left[\overline{(+\text{long})_a} \right] (C_2)_b \quad Y \right]$$

Condition: (1) $(\ )_a$ or $(\quad)_b^{\cdot}$ must be present

Case 1. (Maximum string)

$$V \rightarrow [2 \text{ stress}] \quad / \left[C_0 \left[\overline{+\text{long}} \right] C_2 \quad Y \right]$$

(a) Assign 2 stress to the vowel in the first syllable if it is long and followed by at least two consonants:

hydrosanitation — ① 1 1 — h a ɪ d r o s æ n ɪ r t e ʃ ə n

Case 2. (First subscripted optional string excluded)

$$V \rightarrow [2 \text{ stress}] \quad / \left[C_0 \_\_\_ C_2 \quad Y \right]$$

(a) Assign 2 stress to the vowel in the first syllable if it is followed by at least two consonants:

circumnavigation — 1 1 1 — s ʌ r k ʌ m n æ v ɪ r g e ə n

Case 3. (Second subscripted optional string excluded)

$$V \rightarrow [2\ \text{stress}] \ / \ [\ C_0 \ \left[\begin{smallmatrix} \\ +\text{long} \end{smallmatrix}\right] \ Y\ ]$$

(a) Assign 2 stress to the vowel in the first syllable if it is long:

dielectric	① 1 1 dajɛlɛktrɪk

Algorithmic Application: If the first syllable is strong, i.e., if it contains either a long vowel or two or more consonants assign the vowel primary stress.

Modifications: This rule has been extended to include both the first syllable of the root and the first syllable of the left-most prefix.

This rule has been moved to follow the Compound Stress Rule to prevent the retention of primary stress in prefixes by the Compound Stress Rule in words such as recruit and intend.

Cursory Rule (non-cyclic)

$$V \rightarrow \begin{bmatrix} -\text{long} \\ -\text{stress} \end{bmatrix} / [\ X - \begin{bmatrix} V \\ 1\ \text{stress} \end{bmatrix} \ C_0 \_\_ \ CV \ Y\ ]$$

Condition: (1) if right-most morph is a suffix, check for stress exclusion category.

Algorithmic Application: (only one case of the Cursory Rule) The vowel following the primary-stressed vowel, if it is not the last vowel in the word, is shortened and its stress removed.

Examples: 2 1 2↗° infirmary, 1 2↗° cursory, 1 2↗° curative (/e/→/ɔɛ/, later reduced to /ə/.)

Modifications: Pre-vocalic vowels are not shortened.

The suffixes discussed under the Main Stress Rule which do not affect stress placement are excepted from the domain of the Cursory Rule if they are either word-final or precede another word-final suffix in the same category.

Vowel Reduction Rule (non-cyclic, applicable to all vowels having required context)

$$\begin{bmatrix} \{/\,\varepsilon\,/ \\ /\,\text{ɪ}\,/\} \\ \text{-stress} \end{bmatrix} \rightarrow \bar{\text{ɪ}} \;/\; X\_\_Y$$

$$\begin{bmatrix} V \\ \text{-stress} \\ \text{-long} \end{bmatrix} \rightarrow \text{ə} /\; X\_\_Y$$

Case 1. (reduction of /ɛ/ and /ɪ/)

2 1
ptolemaic (/ɛ/ → /ɪ̄/, /ɪ/ → /ɪ̄/)

Case 2. (reduction of other short non-stressed vowels)

1
curator (/ɚ/ → /ə/, /ɔ/ → /ə/)

Algorithmic Application: All non-long unstressed vowels are reduced, /ɛ/ and /ɪ/ to /ɪ̄/, i.e., reduced I, and all others to /ə/.

Modification: The phonemes /ɛ/ and /ɪ/ are reduced to /ɪ̄/ rather than to /ə/.

A Stress-Dependent Letter-to-Phoneme Rule

The rule which follows appears to be stress-dependent and was placed in the stress placement section rather than with other letter-to-phoneme rules:

Rule: The phoneme /t/ is changed to /č/ and the phoneme /d/ to /ǰ/ if it is not in the initial consonant cluster and precedes unstressed /u/ or /U/, or if it precedes unstressed /ə/ which was /u/ or /U/ before application of stress placement rules.

Example: 2 1 2 1
perpetuity (/t/) perpetual (/č/)

Examples (showing words which do not fit the context of the rule and therefore retain /t/ or /d/ as pronunciation):

1 1 1 1 1
(a) tutor, duty, studious, duration, tureen

In these cases, the /t/ or /d/ is in the initial consonant cluster

2 1 1 1
(b) adumbration, modus, status

The /t/ or /d/ in these cases is not in the initial consonant cluster, but precedes unstressed /ə/ which was not /u/ or /ʊ/ before application of stress rules.

2 1 2 1 1 1 2
(c) institution, centurion, Hindu, constitute

In the above cases, /t/ or /d/ is not in the initial consonant cluster and precedes stressed /u/ or /ʊ/.

The stress program has been modified to effect this change. The phonemes /t/ and /d/ preceding unstressed /u/ or /ʊ/ not in the first syllable are changed following the cyclic rules which place all stress. After the Destressing Rule and the Cursory Rule, a change is also made if the destressed (and possibly shortened) vowel was previously a /u/ or /ʊ/ and not in the first syllable.

A Complete Example

MULTINUCLEOLATED	Input
MULTI=NUCLEOL+ATE+ED	Result of Stage 1
-----=n-kl--l+---+--	Result of Stage 2
mʌlti=nukliol+ɛt+ɪd	Result of Stage 3
1	Main Stress Rule, cycle 1 (domain: multi=nukliol)
1	Stressed Syllable Rule, cycle 1
(1)	Alternating Stress Rule, cycle 1
	Main Stress Rule, cycle 2 (domain: mʌlti=nukliol+ɛt)

	(1)	Stressed Syllable Rule, cycle 2
	(1)	Alternating Stress Rule, cycle 2
		(There are no further cycles since Condition (2) of the MSR applies to +ED)
1		Strong First Syllable Rule
	-stress -long	Destressing Rule
2	1 2	Compound Stress Rule
	ə ɪ	Vowel Reduction Rule
2 1 2 mʌltimuklɪəletɪd		Final Result

IV Reliability

Two studies have been made to determine the accuracy of phonological and stress placement rules and to select a minimal set of rules which will produce accurate results in as many cases as possible. The set of letter-to-phoneme rules used in the first testing procedure contained 534 rules: included were 127 consonant rules, 46 prefix rules (giving pronunciations for 40 prefixes), 155 suffix rules (covering 96 suffixes) and 206 vowel rules. The Trisyllabic Shortening Rule was included in the set of stress rules. A sample of 4,725 words from the Brown Corpus was tested with the following results.

	Number of Words	Percentage given acceptable pronunciation
1- to 5-letter words	2,914	82
7-letter words	1,174	73
12- to 21-letter words	637	65

An acceptable pronunciation is one which is given in Webster's Third International Dictionary,[17] either preferred or alternate. Of the 2,375 1- to 5-letter words which received acceptable pronunciations, 2,135 were given preferred pronunciations, 228 were given alternate pronunciations and 12 received the verbal pronunciation of noun/verb pairs.

A table of frequency of use and statistical accuracy of each rule was derived from this study. These results led to the removal of the Trisyllabic Shortening Rule and to the formulation of eight sets of phonological rules ranging from a maximal set of 557 rules to a minimal set of 277 rules.

In the second study, these eight sets of rules were each applied to a new group of test words which was composed of a random sampling of six-letter words from the Brown Corpus (250 words), the Heritage English Dictionary (150 words) and Stedman's Medical Dictionary (100 words). Results of this study are as follows:

Number of Rules	Percentage given acceptable pronunciation		
	Heritage	Brown Corpus	Stedman's
557	73	69	65
531	72	69	64
453	72	69	64
413	72	67	53
359	70	65	49
308	68	64	43
286	67	63	44
277	66	62	43

Note: The addition of special medical prefixes would increase the accuracy ~~of rules applied to~~ the sample from Stedman's Medical Dictionary by approximately ten per cent.

The set of rules currently being used in the text-to-speech system is the set containing 413 rules A list of the maximal set of 557 rules together with instructions for extracting the other sets of rules is given in the appendix.

There are a number of problem areas, many of which derive from the lack of a lexicon. Problems of this type include incorrect suffix or prefix

recognition and the treatment of compounds as single morphs. Some examples from each problem area are given below:

Mispronunciation of single vowel:

international	/e/	modeled	/ɨ/
menu	/u/	strategically	/ɛ/
environmental	/ɪ/	buried	/yu/
hotels	/a/	two	/o/

The pronunciations of the underlined vowels in the contexts above are encountered infrequently, and, in most cases, are not predictable. In the word international, the context which determines the pronunciation of [a] is the right-hand context [C+iV]. A long vowel almost always is found in this context as in nation, station, explanation, observational, sensational. A short [e] is usually found preceding [C+iC], e.g., malefic, angelic, systemic, photogenic, and is long only in a few words, e.g., strategic, scenic, and in the suffix +plegic. There are very few words ending in the vowel [u]--most are either low frequency words or proper names. The palatalization in menu is not found in other words with final [u], e.g., flu, emu, gnu, impromptu. The word two is very irregular in pronunciation. Most words ending in [o] such as go, no, so, calico, echo have the sound /o/. It may be noted, however, that two other words which, like two, are very high frequency words, have the same pronunciation of final [o] as two, i.e., do, to. The mispronunciation of the [e] in modeled is due to the assumption, lacking a lexicon, that the morphemic analysis is model + ed.

Mispronunciation of vowel digraph:

said	/e/	shoes	/o/
break	/i/	guitars	/u/
forfeit	/i/	should, would	/o/
endowed	/o/	theirs	/i/

The reasons for mispronunciation of the vowel digraphs underlined above fall into a number of categories. There are very high frequency words, said, should and would, which do not follow letter-to-sound rules. Said may be contrasted with the words laid, maid, paid, and raid; the words should and would contrast with mould, shoulder and boulder. The sequence [eir] as in theirs, heir, weir is not found frequently in English, nor is the sequence [feit] as in forfeit, surfeit and counterfeit. Rules for [ei] in these two contexts were considered unproductive. Final [oe] in English is usually pronounced as in oboe, toe and foe; the pronunciation found in shoe, and also in canoe, is rare. Rules governing the pronunciation of [ow] (endowed) and [ui] (guitars) are statistically based. Although there are ~~many words in which similar non-context dependent pronunciations are found~~, e.g., cow, allow, eyebrow and build, guilt, guinea, other pronunciations are statistically more likely, e.g., those found in shadow, glow, follow and bruise, juice, nuisance. The pronunciation of break is not predictable -- the word steak has the same digraph pronunciation, but other similar words such as creak, freak and streak are pronounced like the majority of words containing the digraph [ea].

Mispronunciation of single consonant:

of	/f/	corp	/p/
eager	/j/	exhaust	/h/
two	/w/	physiological	/s/
deserts	/s/	schizophrenic	/z/

The consonants underlined above are either silent or have unpredictable or unusual pronunciations. Silent consonants are found in two, corp and exhaust. The word two is a high-frequency word in which both the [w] and [o] have unusual pronunciations. Silent [w] is rare, although it is also found in the word sword. Final silent [p], as found in corp is also rare. (This word is considered in this section because the pronunciation of both the [r] and the [p] are determined by rules for single consonants.) There are a few words, like exhaust, in which [h] is silent following [ex], e.g. exhibit, exhilarate, exhort and exhume. However, this rule is not sufficiently productive to merit inclusion.

The letter [g] preceding [e], [i] and [y] in English usually has a soft sound as in integer and wager. In particular, many words ending in [ger] are a combination of a root with final [e] and the suffix [+er], e.g., forager, manager, merger, all of which have a soft [g] sound. The pronunciation of the [g] in eager is unusual and not predictable. Another pronunciation which is frequently unpredictable, i.e., not context-dependent, is that of the letter [s] between vowels. The rule for this context predicts the more frequent sound /s/ whereas the sound /z/ is found in deserts and physiological. The letter [z] in schizophrenic has the rare pronunciation /ts/, and the word if,

as previously discussed, is the only English word in which a final [f] is pronounced as /v/.

Mispronunciation of consonant cluster:

chef	/c/	laugh	/-/
would, should	/ld/	cliches	/č/
calf	/lf/	issue	/s/
tsar	/ts/	these	/ð/

Consonant clusters are infrequently mispronounced The cluster [ch] is the most frequent problem in this category, its pronunciation being determined, in many cases, by the Greek or Latin origin of the word in which it appears. The pronunciation /ʃ/ as in chef and cliches is less frequent than either /č/, e.g., church, or /k/, e.g., chemical. Morph-final [gh] may be pronounced either /f/ as in laugh, enough and cough or, with slightly higher probability, not pronounced, as in high, weigh and dough. Unusual and rare pronunciations of the clusters ts, ss, ld and lf are found in other words above. The pronunciation of [ss] as /ʃ/ is found preceding certain suffixes, e.g., depression, fissure, but rarely within a morph, (tissue, above). Russian orthography is still reflected in the English spelling of tsar even though the pronunciation has been Anglicized. A silent [l] appears in could, would and calf. The words could and would are high frequency words and also differ from regular pronunciation in the vowel digraph [ou]. Although half like calf, also has a silent [l], in most words a final [lf] is pronounced /lf/, e.g., elf, shelf, self, gulf. The high-frequency-word pronunciation of morph-initial [th] as /ð/, e.g., these, then, the, has been discussed previousl

Incorrect suffix recognition:

water	/wɛt+ɚ/	guy	/gyu+i/
thus	/θʌ+z/	disengagement	/dɪsɛŋ+ɪ̄j+mɪ̄nt/
relying	/rɪl+i+ɪ̄ŋ/	exist	/ɛgzɪst
heated	/hi+et+ɪ̄d/	alas	/ələ+z/

Almost all problems in this category arise from the lack of a morph lexicon. Words are pronounced incorrectly because letter strings in a root which appear to be suffixes are converted to phonemes using rules for suffix pronunciation. It may be seen in the examples above that a mistake in morph analysis can cause obvious errors in pronunciation.

Incorrect prefix recognition:

unit	/ən=ɪ̄t/	decimal	/dɪ̄=sɛm+əl/
cool	/ko=al/	emerald	/ɛm=ɛrʌld/
realm	/ri=am/	encouragement	/ɛn=ko=ʌr+ɪ̄j+mɪ̄nt
deem	/di=ɛm/	experimenters	/ɛks=pər=ɪ̄mət+ər+z/

Mistakes in morph analysis produce prohunciation errors, as in the previous category.

Non-recognition of prefix

apart	/əpart/	thermoinhibitory	/θɚmoɪnhɪbətɔri/
refer	/rɪf+ər/	berated	/bɛr+et+ɪd/
dissent	/dɪs+ənt/	correspondent	/ko=rɛspand+ənt/
hydrocele	/hɪdrəsil/	pericardiorrhaphy	/pɛ=ɪkɑrdɪɔrəf+i/

There are many technical prefixes which have not been included in the prefix list. These may be added by a user with particular technical needs. A few prefixes such as [a] in apart and [e] in eject have not been

included because a high error rate would result, i.e., all words beginning in a or e would be incorrectly analyzed. In the remaining cases, prefixes were incorrectly analyzed as part of a root after suffixes were incorrectly removed. Errors in pronunciation, and particularly in stress are the result.

Incorrect stress:

Most of the words in this category have unusual stress patterns which are unpredictable. A comparison with similar words shows the regular stress pattern:

1 motel	1 1 1 morsel, gavel, fuel
1 palette	1 1 1 majorette, curette, suffragette
1 sonata	1 1 1 vertebra, camera, automata
1 urea	1 1 1 trachea, azalea, miscelanea
1 uncomfortable	1 1 unaffordable, uncontestable
1 lunatic	1 1 1 erratic, fanatic, aromatic
1 renegade	1 1 1 brigade, serenade, marinade

The word selects is stressed incorrectly due to lack of information concerning its part of speech (c.f., discussion of modifications in the Main Stress Rule).

The results of this study indicate that the letter-to-phoneme system is quite powerful, even in isolation. When considered in the domain of the over-all text-to-speech system in which a lexicon is available for high-frequency words and compounds, the letter-to-phoneme system should be highly reliable.

NOTES

1. Allen, J. "Speech Synthesis from Unrestricted Text" in Speech Synthesis, edited by J. L. Flanagan and L. R. Rabiner, Stroudsburg, 1973.

2. Allen, J. "Reading Machines for the Blind: The Technical Problems and Methods Adopted for their Solution," IEEE Transactions on Audio and Electroacoustics, Vol. Au-21, No. 3, June 1973, 259-264.

3. Allen, J. "Synthesis of Speech from Unrestricted Text," Proceedings of the IEEE, Vol. 64, No. 4, April 1976, 433-442.

4. Brackets will be used to denote letter, slashes to denote phonemes.

5. Hunnicutt, S. "Pronunciation of High-Frequency English Words," Natural Language Processing Group Memo, M.I.T., 1973.

6. Kucera, H. and Francis, W. N. Computational Analysis of Present-Day American English, Providence, R. I., Brown University Press, 1967.

7. The symbol C is used to denote a consonant, the symbol V to denote a vowel.

8. Venezky, R. A Study of English Spelling-to-Sound Correspondences on Historical Principles, Ann Arbor, 1965.

9. This source is available on tape from J. Olney, System Development Corporation, Santa Monica, California.

10. Walker, J. Rhyming Dictionary of the English Language, London, 1924

11. The American Heritage Dictionary of the English Language, Paperback edition, New York, 1973.

12. Stedman's Medical Dictionary, Baltimore, 1961

13. Although some phonologists would consider both prevocalic [i] and morph-final [o] to be underlyingly short (lax), and would apply a lengthening (tensing) rule to them after stress placement (N. Chomsky and M. Halle, The Sound Pattern of English (New York, 1968), p. 74; M. Halle and S.J. Keyser, English Stress (New York, 1971) p. 30), it has been necessary, because of the difficulty of preserving orthographic information, to consider both vowels as being long before stress placement, exempting them from stress by a contextual analysis. The replacement of the required feature [-low] in the lengthening rule for morph-final vowels by the feature [-constricted pharynx] is due to Halle (Course 23.761, M.I.T.), Nov. 1972.

14. Class lectures, Course 23.761, M.I.T., Fall 1972. Professor Halle has since revised this set of rules.

15. Changes to the stress rules due to the incorporation of a parsing algorithm in the system will be made during the summer, 1976.

16. Hunnicutt, S. "Removal of the Trisyllabic Shortening Rule from Stress," Natural Language Processing Group Memo, M.I.T., 1975.

17. Webster's Third New International Dictionary, Springfield, 1966.

APPENDIX

The set of rules which follows is a maximal set. The n^{th} largest set of rules may be derived by deleting all lines preceded by integers less than n.

Preceding the rules are three lists of variables. (A variable is symbolized by a letter or the sequence "up-arrow" - letter, e.g., ↑L preceded by a dollar sign.) The first list contains only two variables, $B, which represents either a word boundary (#) or a suffix boundary (+), and $D, which represents either a word boundary or a prefix boundary (=). The second list consists of variables representing letters, and the final list, of variables representing phonemes A table of the symbols used to represent phonemes with the corresponding IPA symbol immediately precedes the list of rules:

The general form of a rule is

$$S_1 > S_2 \quad / \quad S_3 \leftarrow S_4$$

where the S_i are strings ($S_1 \neq \emptyset$), / (slash) is one of four possible contexts to be explained below, and ← (back-arrow) is a place-holder for S_1. The string S_1 is a letter string which is to be converted into the phoneme string S_2 if it (S_1) is preceded by S_3 and followed by S_4. S_3 and S_4 may be either letter strings or phoneme strings set by the following context markers:

character	expects	left context	right context
]		letter	letter
[		phoneme	letter

continued:

character	expects	left context	right context
)		letter	phoneme
(		phoneme	phoneme

In a left context, variables are followed, rather than preceded by a dollar sign.

The final two sections in the set of rules are marked PREFIX and SUFFIX respectively. Prefixes are listed left-to-right longest-match-first fashion (e.g., extra precedes ex, uni precedes un) within a left-to-right alphabetical sort.

Suffixes are listed right-to-left longest-match-first, (e.g., uous precedes ous which precedes us which precedes s) within a right-to-left alphabetical sort. Each suffix is preceded by a letter and a 2- or 3-digit number. The letter represents the type of suffix as follows:

P	a non-vocalic suffix (suffix which does not begin with a vowel) before which a [y] may change to [i]
E	a suffix (generally vocalic and very short) which is recognized only before inflectional and consonantal suffixes
F	regular suffix

The number preceding each suffix is the octal representation of a code describing the syntactic compatibility of that suffix with other suffixes. The right-most digit represents the possible parts of speech of a preceding suffix (a suffix to the left of the suffix under consideration). If the suffix may be preceded by an adjectival suffix, the digit is

1, if possibly preceded by a verbal or nominal suffix, the digit is 2 or 4 respectively; if the suffix may be preceded by a combination of these, the digit is the sum of the corresponding digits.

The digit immediately to the left of the right-most digit represents the possible parts of speech of the suffix under consideration; it is calculated as above. If the number contains 3 digits, its left-most digit is a 1, indicating that the suffix occurs only at the end of a word.

Example: F 166-ES

The suffix is a regular (non-special) suffix; it occurs only in word-final position; it forms nouns (4) and verbs (2) and may be preceded by a nominal (4) or verbal (2) suffix. (In each case, 4 and 2 are summed.)

PHONEME TABLE

Rules.	IPA	Rules	IPA
E	i	R	r
↑I	I	L	l
A	eɪ	J	y
↑E	ɛ	H	h
"A	æ	P	p
↑A	a	B	b
↑O	ɔ	M	m
O	o	T	t
↑U	ʊ	D	d
U	u	N	n
↑Y	ʌ	K	k
YR,↑	ɝ	G	g
I	aɪ	↑G	ŋ
"O	ɔɪ	F	f
"U	aʊ	V	v
"I	ɨ,ɪ	↑T	θ
Y	ə	↑D	ð
↑L	.l	S	s
↑R	.r	Z	z
YL	·l	↑S	ʃ
YR	·r	↑Z	ʒ
W	w	↑C	ʧ
↑W	ʍ	↑J	ʤ

```
/$^P IS PART OF $^N

$B, #, +
$D, #, =

/LETTER VARIABLES
$W, A, E, I, O, U, Y, +A, +E, +I, +O, +U, +Y, LE#, RE#
$F, I, E, Y, +I, +E, +Y
$^F, A, O, U, +A, +O, +U
$L, B, C, D, F, G, J, K, L, M, N, P, R, S, T, V, Z, X, +C, +D, +F, +H, +L  +M, +N, +P, +R, +S, +T, +W, +'S
$X, +IENT, +IOUS, +ION, +IAL, +IAN, +IENCE, +EON, IA#
$^A, P, T, K, F, S, C, SH, CH, H, +T, +S
$^B, A, E, I, O, U, Y, W, H, B, G, L, M, N, R, V, Z
$^I, +ED, +ING, +ER, +EST, +ES, E#, E+
$^E, N, L, LL
$^K, +ES, +ED
$S, A, E, I, O, U, Y, R, P, K, C
$V, N, S
$T, L, S
$R, A, E, I, O, U, Y, R
$^L, +IC, +ITY, +IFY, +IENT, +IOUS, +ION, +IAL, +IAN, +IENCE
$^M, +IVE, +AL, +OUS, +AGE, +LY
$^N, +IBLE, +IBLY, +IBILITY, +IFIC, +IC, +ACE, +ICE, +URE, +UTE, +ULE, +ITY, +IFY, +ATIVE
$^P, +ATORY, +ITION, +IFICATIVE, +IFICATION, +OON, +ARCHY, +ULAR, +UAL, +UOUS
$^O, +SIS, +TY
$E, TH, ST
$H, L, R


/PHONEME VARIABLES
$C, P, T, K, B, D, G, F, S, V, Z, M, N, L, R, W, J, H, ^T, ^D, ^S, 'Z, 'C, ^J, 'G
$G, P, B, F, M
$J, P, T, K, F, ^T, S, ^S, ^C, H
$K, A, E, I, O, U, Y, W, "A, ^A, ^I, ^O, "O, ^U, "U, ^Y, "I, H, B, G, ^J, L, M, N, ^G, R, ^D, V, Z, ^Z, 'L,
^M, D, ^R
$^H, K, G
$M, D, T
$N, S, Z, ^S, ^C, ^Z, ^J
$Q, A, ^A, "A, E, ^E, I, ^I, O, ^O, "O, U, ^U, "U, Y, ^Y, "I, +A, +^A, +"A, +E, +^E, +I, +^I, +O, +'O,
+"O, +U, +^U, +"U, +Y, +^Y, +"I
$Z, #, =, B, P, M, F, V, K, G, H, T, D, N
$I, ^J, ^C, ^S, ^Z, J
$A, ^D, ^Z, V


    %QCON

4   CHEM   > K^EM   ] <
6   PSYCH  > SIK    ] <
3   TECH   > T^EK   ] <
6   ARCH   > ^ARK   ] <
```

```
   /TRIPLE CONSONANTS

2  CCH > K     ] ←
5  CHR > KR    ] ←
6  CHL > KL    ] ←
2  CHN > KN    ] ←
   NCH > N^C   ] ← $B
6  NCH > ^GK   ] ←
2  NGU > ^G    ] ← E$B
5  NQU > ^GKW  ] ←
5  SCH > SK    ] ←
2  THM > ^D'M  ] ← +$L
2  THM > ^D'M  ] ← #
   TCH > ^C    ] ←

   /DOUBLE CONSONANTS

   SH > ^S     ]       ←
   TH > ^D     ]       ← $
   TH > ^T     ]       ←
4  BT > TT     ]       ← #
4  BT > TT     ]       ← $^I
6  CC > KS     ]       ← $F
   CC > KK     ]       ←
   CK > KK     ]       ←
   CH > ^C     ]       ←
   DG > ^J^J   ]       ←
5  DJ > ^J^J   ]       ←
4  GG > GG     ]       ← $B
4  GG > ^J^J   ]       ← $F
3  GH > G      ] D$    ←
   GH >        ]       ←
4  GN > N      ] D$    ←
3  GN > N      ]       ← #
4  GN > N      ]       ← +$L
4  GN > N      ]       ← $'I
2  GM > M      ]       ← $B
5  GU > G      ]       ← $^I
2  GU > GW     ] N     ← $W
   GU > GUU    ]       ← $B
2  GU > GW     ] W$    ← $W
   KN > N      ] D$    ←
4  LK > KK     ]       ← #
4  LK > KK     ]       ← $^I
4  LK > KK     ]       ← +$L
4  LM > M      ] A     ← $B
4  MB > MM     ]       ← #
4  MB > MM     ]       ← $^I
4  MB > MM     ]       ← +$L
4  MN > MM     ]       ← #
4  MN > MM     ]       ← $^I
4  MN > MM     ]       ← +$L
   NK > ^GK    ]       ←
2  NC > N^C    ]       ← $X
   NC > NS     ]       ← $F
   NC > ^GK    ]       ←
   NG > 'G     ]       ← #
7  NG > ^G     ]       ← +$L
4  NG > N^J    ]       ← +ES#
   NG > N^J    ]       ← E#
7  NG > ^G     ]       ← $^I
   NG > 'GG    ]       ←
3  NX > ^GKS   ]       ←
   PH > F      ]       ←
   QU > K      ]       ← $^I
3  QU > K      ]       ← ET$B
   QU > KW     ]       ←
2  RH > R      ]       ←
3  ST > S      ]       ← LE$B
4  ST > S      ] W$    ← +EN
2  SC > ^S     ] S$    ← $X
6  SC > 'C     ] V$    ← $X
   SC > SS     ]       ← $F
2  SC > SS     ]       ← LE$B.
   SC > SK     ]       ←
2  SM > Z'M    ]       ← +$L
2  SM > Z^M    ]       ← #
   SS > ^S     ]       ← $X
3  SS > ^S     ]       ← +URE
2  TZ > TS     ]       ←
   WH > HW     ] D$    ←
   WR > R      ] D$    ←

   /SINGLE CONSONANTS

   B > B       ]       ←
6  C > ^S      ] S$    ← $X
2  C > ^C      ] V$    ← $X
3  C > ^S      ] S$    ← +EOUS
3  C > ^C      ] V$    ← +EOUS
   C > S       ]       ← $F
   C > K       ]       ←
   D > D       ]       ←
   F > F       ]       ←
   G > ^J      ]       ← $F
   G > G       ]       ←
   H >         ]       ← $B
3  H >         ] W$    ← $L
   H > H       ]       ←
   J > ^J      ]       ←
   K > K       ]       ←
   L > L       ] H$    ←
```

```
    L > ^L   ] L$ ← E#
    L > ^L   ] L$ ← $^K
    L > L    ]    ←
    M > M    ]    ←
    N > N    ]    ←
    P > P    ] D$ ← $H
    P >      ] D$ ← $L
    P > P    ]    ←
    Q > K    ]    ←
    R > R    ] R  ←
    R > ^YR  ] L$ ← E#
    R > ^YR  ] L$ ← $^K$B
    R > R    ]    ←
2   S > ^S   ] T$ ← $X
3   S > ^S   ] T$ ← U+
2   S > ^S   ] T$ ← +U
5   S > ^C   ] N  ← $X
    S > ^C   ] N  ← U+
3   S > ^C   ] N  ← +U
5   S > ^Z   ] R$ ← $X
    S > ^Z   ] R$ ← U+
2   S > ^Z   ] R$ ← +U
    S > S    ] D$ ←
    S > S    ] W$ ← $^
    S > Z    ] W$ ← $L
    S > Z    ] W$ ← E$B
1   S > Z    ] M  ← $^B
    S > Z    ] W$ ← +$W
    S > S    ]    ←
    T > ^S   ] S$ ← $X
2   T > ^C   ] V$ ← $X
    T > T    ]    ←
    V > V    ]    ←
    W > W    ] D$ ←
    W > W    [ C$ ←
2   X > Z    ] D$ ←
    X > KS   ]    ←
    Y > J    ] D$ ←
    Z > Z    ]    ←

    %PREFIX PHO

5   ANTI=    > "ANTI=   ( ←
    BE=      > B^I=     ( ←
    BI=      > BI=      ( ←
5   CIRCUM=  > SYRKYM=  ( ←
    COM=     > K^AM=    ( ←
    CON=     > K^AN=    ( ←
    CO=      > K^A=     ] ← $L$L
    CO=      > KO=      ( ←
    DE=      > DE=      ] ← $W
    DE=      > D^I=     ( ←
    DIS=     > D^IS=    ( ←
    EM=      > ^IM=     ( ←
    EN=      > ^IN=     ( ←
    EXTRA=   > ^EKSTRY= ( ←
    EX=      > ^IGZ=    ] ← $W
    EX=      > ^IKS=    ( ←
4   HYPER=   > HIPYR=   ( ←
    IM=      > ^IM=     ( ←
    INTER=   > ^INTYR=  ( ←
4   INTRA=   > ^INTRY=  ( ←
    IN=      > ^IN=     ( ←
6   MAL=     > M"AL=    ( ←
1   MAXI=    > M"AKS^I= ( ←
5   MICRO=   > MIKRO=   ( ←
1   MINI=    > M^INE=   ( ←
    MIS=     > M^IS=    ( ←
2   MULTI=   > MYLT^I=  ( ←
1   NEO=     > NEO=     ( ←
    NON=     > N^AN=    ( ←
    PER=     > PYR=     ( ←
    PRE=     > PRE=     ( ←
    PRO=     > PRO=     ( ←
2   PSEUDO=  > SUDO=    ( ←
5   PSYCHO=  > SIKO=    ( ←
    RE=      > RE=      ] ← $L
    RE=      > R^I=     ( ←
3   SELF=    > S^ELF=   ( ←
2   SEMI=    > S^EME=   ( ←
4   SUB=     > SYB=     ( ←
3   SUPER=   > SUPYR=   ( ←
    TRANS=   > TR"ANS=  ( ← $J
2   TRANS=   > TR"ANS=  ] ← $W
    TRANS=   > TR"ANZ=  ( ←
4   ULTRA=   > YLTRY=   ( ←
    UNDER=   > YNDYR=   ( ←
4   UNI=     > JUN^I=   ( ←
    UN=      > YN=      ( ←

    /SUFFIX PHO

    +ABILITY > +YB^IL^ITE ( ←
    +ABLE    > +YBYL      ( ←
5   +ABLY    > +YBLE      ( ←
2   +ACE     > +A^S       ] ← $X
    +ACE     > +^YS       ( ←
2   +AGE     > +A^J       ] ← $^M
    +AGE     > +^I^J      ( ←
    +AL      > +AL        ] ← $X
    +AL      > +"AL       ( ←
    +ANCE    > +"ANS      ( ←
```

	+ANT	>	+"AN^C	]		← $X	|		+IAL	>	+E"AL	(		←
	+ANT	>	+"ANT	(		←	|	2	+IAN	>	+EAN	]		← $X
4	+AN	>	+AN	]		← $X	|	2	+IAN	>	+E"AN	]		← $^L
	+AN	>	+"AN	(		←	|	3	+IAN	>	+YN	(	I$	←
3	+ARCHY	>	+^ARKE	(		←	|		+IAN	>	+E"AN	(		←
4	+ARIAN	>	+^EREYN	(		←	|		+IBILITY	>	+^IB^IL^ITE	(		←
4	+ARY	>	+^ER^I	]		← +LY	|		+IBLE	>	+'IBYL	(		←
	+ARY	>	+^ERE	(		←	|	5	+IBLY	>	+^IBLE	(		←
	+AR'	>	+"AR	(		←	|	4	+ICE	>	+"IS	(		← #
	+ATE	>	+YT	]		← $^M	|	5	+IC	>	+'I^S	]		← $X
	+ATE	>	+A'S	]		← $X	|	5	+IC	>	+^IS	]		← $F
7	+ATE	>	+"AT	]		← $'L	|		+IC	>	+'IK	(		←
	+ATE	>	+AT	(		←	|	2	+IENCE	>	+YNS	(	I$	←
5	+ATIVE	>	+AT^IV	(		←	|		+IENCE	>	+E^ENS	(		←
4	+ATORY	>	+YT^ORE	(		←	|	3	+IENT	>	+E^EN^C	]		← $X
5	+ATRY	>	+"ATRE	(		←	|	2	+IENT	>	+E^ENT	]		← $'L
3	+CRACY	>	+KR^ASE	(		←	|	2	+IENT	>	+YNT	(	I$	←
1	+DOM	>	+D^AM	(		←	|		+IENT	>	+E'ENT	(		←
5	+EAL	>	+E"AL	(		←	|	3	+IETY	>	+I^ETE	(		←
	+ED	>	+^ID	]		← +$L	|		+IFICATION	>	+^IF^IKA^SYN	(		←
	+ED	>	+^ID	(	M$ ←		|	2	+IFICATIVE	>	+^IF^IKAT^IV	(		←
	+ED	>	+D	(	K$ ←		|		+IFIC	>	+'IF'IS	]		← $F
	+ED	>	+T	(		←	|		+IFIC	>	+^IF^IK	(		←
4	+EER	>	+^IR	(		←	|	5	+IFY	>	+'IFI	(		←
6	+EE	>	+EE	(		←	|	5	+ILE	>	+'IL	(		←
7	+ENCE	>	'+^ENS	(		←	|	1	+IMONY	>	+^IMONE	(		←
	+ENT	>	+^EN^C	]		← $X	|		+ING	>	+^I^G	(		←
	+ENT	>	+^ENT	(		←	|	2	+ION	>	+EON	]		← $X
	+EN	>	+^EN	(		←	|	2	+ION	>	+E^AN	]		← $'L
5	+EON	>	+YN	(	I$ ←		|		+ION	>	+YN	(	I$	←
5	+EON	>	+E^AN	(		←	|		+ION	>	+E^AN	(		←
5	+EOUS	>	+YS	]	C ←		|	7	+IOUS	>	+YS	(	I$	←
5	+EOUS	>	+EYS	(		←	|	7	+IOUS	>	+EYS	(		←
5	+ERY	>	+^ERE	(		←	|		+ISH	>	+^I^S	(		←
	+ER	>	^ER	]		← +AT	|	2	+ISM	>	+'IZYM	]		← +$L
	+ER	>	+YR	(		←	|		+ISM	>	+^IZYM	(		← #
2	+ESCE	>	+^ES	(		←	|		+ISM	>	+^IZM	(		←
	+EST	>	+'EST	(		←	|		+IST	>	+^IST	(		←
2	+ESQUE	>	+^ESK	(		←	|	5	+IS	>	+"IS	(		←
	+ES	>	+^IZ	(	N$ ←		|		+ITION	>	+^I^SYN	(		←
	+ES	>	+Z	(	K$ ←		|	2	+ITIS	>	+ITYS	(		←
	+ES	>	+S	(		←	|		+ITY	>	+^ITE	(		←
3	+ETTE	>	+^ET	(		←	|		+IVE	>	+^IV	(		←
1	+ETH	>	+^E^T	(		←	|		+IZE	>	+IZ	(		←
	+EY	>	+E	(		←	|	4	+LESS	>	+L^ES	(		←
1	+FOLD	>	+FOLD	(		←	|	4	+LET	>	+L'ET	(		←
5	+FUL	>	+FYL	(		←	|	3	+LIKE	>	+LIK	(		←
6	+HOOD	>	+H^UD	(		←	|		+LOG	>	+L^A^J	]		← $F
2	+IAL	>	+EAL	]		← $X	|		+LOG	>	+L^AG	(		←
3	+IAL	>	+E"AL	]		← $'L	|		+LY	>	+LE	(		←
	+IAL	>	+YL	(	I$ ←		|		+MENT	>	+M^ENT	(		←

```
2  +METER    > +M^ETYR   (     ←        |      +Y       > +^I        ]     ← +LY
2  +METRY    > +M^ETRE   (     ←        |      +Y       > +E         (     ←
1  +MOST     > +MOST     (     ←        |      +'S      > +^IZ       ( N$  ←
   +NESS     > +N^ES     (     ←        |      +'S      > +S         ( J$  ←
4  +OGIZE    > +^O^JIZ   (     ←        |      +'S      > +Z         (     ←
   +OID      > +"OD      (     ←        |
3  +OLIZE    > +^OLIZ    (     ←        |    /DIGRAPH RULES: A-
4  +OMETER   > +^AM^ETYR (     ←        |
2  +OMETRY   > +^AM^ETRE (     ←        |  4  AE > ^E    ( D$       ←
2  +OMIZE    > +^OMIZ    (     ←        |  4  AE > E     (          ←
5  +OON      > +UN       (     ←        |     AI > "AY   ]          ← R
2  +OPHIZE   > +^OFIZ    (     ←        |     AI > A     (          ←
1  +OQUIZE   > +^OKWIZ   (     ←        |     AU > ^O    (          ←
1  +OQUY     > +^OKWE    (     ←        |     AY > A     (          ←
2  +ORY      > +^OR^I    ]     ← +LY    |     AW > ^O    (          ←
   +ORY      > +^ORE     (     ←        |
   +OR       > +^OR      (     ←        |     /DIGRAPH RULES: E-
   +OSIS     > +OSYS     (     ←        |
   +OUS      > +YS       (     ←        |  4  EA > E     ( D$C$C$ ← #
5  +PHILE    > +FIL      (     ←        |  5  EA > E     ( D$C$   ← #
1  +PLASTY   > +PL"ASTE  (     ←        |  5  EA > EY    (        ← #
2  +RY       > +R^I      ]     ← +LY    |     EA > ^Y    (        ← R$C
   +RY       > +RE       (     ←        |  6  EA > E     (        ← R
   +S'       > +^IZ      ( N$  ←        |  4  EA > ^E    (        ← L$C
   +S'       > +S        ( J$  ←        |  2  EA > E     (        ← $A+$Q
   +S'       > +Z        (     ←        |  2  EA > ^E    (        ← $A$Q
3  +SELF     > +S^ELF    (     ←        |     EA > E     (        ←
   +SHIP     > +^S^IP    (     ←        |
6  +SIS      > +S^IS     (     ←        |  2  EI > -I    ] D$S    ←
3  +SOME     > +SYM      (     ←        |  6  EI > A     ]        ← G
   +S        >           ] S   ←        |  4  EI > A     ] V      ←
   +S        > +S        ( J$  ←        |     EI > E     (        ←
   +S        > +Z        (     ←        |  1  EU > J^U   [ D$C$Z$ ← R
   +TY       > +TE       (     ←        |  3  EU > J^U   [ D$Z$   ← R
5  +UAL      > +J^UWYL   (     ←        |  1  EU > ^U    [ D$C$C$ ← R
3  +ULAR     > +J^ULYR   (     ←        |  2  EU > ^U    [ D$C$   ← R
   +ULE      > +JUL      [ Z$  ← #      |  4  EU > Y     ]        ← R$B
   +ULE      > +UL       (     ← #      |  2  EU > JU    ( D$C$Z$ ←
5  +ULE      > +JUL      [ Z$  ← $^I    |  4  EU > JU    ( D$Z$   ←
5  +ULE      > +UL       (     ← $^I    |     EU > JU    ( D$     ←
   +ULE      > +J^UL     [ Z$  ←        |  2  EU > U     ( D$C$C$ ←
   +ULE      > +^UL      (     ←        |  2  EU > U     ( D$C$   ←
3  +UOUS     > +J^UWYS   (     ←        |  1  EU > JU    [ Z$     ← $L$^L
   +URE      > +J^UR     (     ←        |  1  EU > U     ]        ← $L$^L
5  +US       > +^YS      (     ←        |  1  EU > EY    ]        ← $L$B
5  +UTE      > +JU^S     [ Z$  ← $X     |     EU > JU    ( Z$     ←
5  +UTE      > +JUT      [ Z$  ←        |     EU > U     (        ←
5  +UTE      > +U^S      ]     ← $X     |     EW > JU    ( Z$     ←
5  +UTE      > +UT       (     ←        |     EW > U     (        ←
2  +WISE     > +WIZ      (     ←        |     EY > A     (        ←
3  +WORTHY   > +WYR^DE   (     ←        |
```

```
   /DIGRAPH RULES  I-

2  IA  > I"A   ( D$C$C$ ←
   IA  > I"A   ( D$C$,  ←
1  IA  > IY    ]       ← H$B
2  IA  > JY    ] W$^E$ ← $B
3  IA  > EY    (       ← $B

3  IEU > JU    (       ←
7  IE  > I     (       ← #
2  IE  > I     (       ← +$C
   IE  > E     (       ←


4  IU  > EY    (       ← $C$B

   /DIGRAPH RULES. O-

4  OA  > OY    (       ← #
   OA  > O     (       ←
   OE  > O     (       ← #
2  OE  >~Y     ]       ← R
   OE  > O^E   (       ←
   OI  > "O    (       ←
6  OU  > U     (       ← $B
4  OU  > ^O    ]       ← GHT
5  OU  > U     ]       ← P
   OU  > O     (       ← L$C
2  OU  > "U    ( D$C$C$ ← R$B
   OU  > "U    ( D$C$   ← R$B
3  OU  > O     ( C$     ← R$C$B
2  OU  > O     ] L$     ← R$LE$B
   OU  > Y     ]       ← R
   OU  > "U    (       ←

3  OY  > "O    (       ←
   OW  > O     (       ←

   /DIGRAPH RULES  U-

1  UA  > ^A    ]       ← $L+IC
1  UA  > WA    ] S     ←
   UE  > U     (       ← $B
7  UE  > JU^E  ( Z$    ←
3  UE  > U^E   (       ←
1  UI  > ^I    ( K     ← T
   UI  > U     (       ←
```

```
   /DIGRAPH RULES. DOUBLE VOWELS

1  AA  > ^A    (       ←
2  EE  > ^I    ]       ← R
   EE  > E     (       ←
1  II  > E     (       ←
4  OO  > ^U    ]       ← R
7  OO  > U     ( SC$   ←
   OO  > ^U    (       ←
   OO  > U     (       ←
1  UU  > JUY   ( Z$    ←
1  UU  > UY    (       ←

   /VOWEL RULES

   A > ^A   ]          ← #
4  A > ^A   ]          ← LM#
4  A > ^O   ]          ← LK#
   A > ^O   ]          ← LL$^I
   A > "A   ]          ← X
   A > ^O   ( W        ← R$C
1  A > "A   ]          ← R$W
   A > ^A   ]          ← RR$^I
   A > "A   ]          ← RR
   A > ^A   ]          ← R
   A > A    ]          ← $EE$B
6  A > A    ]          ← $LI$W
6  A > A    ]          ← $L+I$W
6  A > A    ]          ← $LE$W
6  A > A    ]          ← $L+E$W
   A > A    ]          ← $L$^I
5  A > "A   ( W        ← $^H
5  A > "A   [ D$C$C$   ← $L$^N
2  A > "A   [ D$C$     ← $L$^N
2  A > "A   [ D$       ← $L$^N
   A > A    [ D$C$C$   ← $L+$W
   A > A    [ D$C$     ← $L+$W
   A > A    [ D$       ← $L+$W
4  A > A    [ D$C$     ← $L$HE$B
2  A > A    [ D$C$     ← $L$H$^I
   A > A    ]          ← $W
   A > ^A   ( W        ←
   A > "A   (          ←
5  E > E    ( D$C$C$   ← $B
   E > E    ( D$C$     ← $B
   E >      (          ← #
5  E >      [ ^H$      ← +$^F
6  E > ^E   ]          ← $'Q
   E >      ]          ← +$L
   E > ^E   ]          ← X
   E > ^E   (          ← L$C
```

```
   E > ^E  ]          ← RR
   E > ^Y  (          ← R$C
   E > ^I  ]          ← RE$B
5  E > E   ]          ← $LI$W
5  E > E   ]          ← $L+I$W
5  E > E   ]          ← $LE$W
5  E > E   ]          ← $L+E$W
   E > E   ]          ← $L$^I
2  E > ^E  [ D$C$C$   ← $L$^N
2  E > ^E  [ D$C$     ← $L$^N
3  E > ^E  [ D$       ← $L$^N
   E > E   [ D$C$C$   ← $L+$W
   E > E   [ D$C$     ← $L+$W
   E > E   [ D$       ← $L+$W
   E > E   ]          ← $W
   E > ^E  (          ←
   I > E   (          ← #
   I > ^I  ]          ← X
   I > ^Y  (          ← R$C
2  I > I   ]          ← $EE$B
5  I > ^I  ]          ← $L+I$W
5  I > ^I  ]          ← $L+E$W
   I > I   ]          ← $L$^I
3  I > ^I  [ D$C$C$   ← $L$^N
4  I > ^I  [ D$C$     ← $L$^N
   # > I   ]          ← GH
3  I > I   ]          ← ND$B
   I > I   [ D$C$C$   ← $L+$W
   I > I   [ D$C$     ← $L+$W
   I > I   [ D$       ← $L+$W
   I > ^I  ]          ← LL
5  I > I   [ D$C$     ← $L$HE$B
2  I > I   [ D$C$     ← $L$H$^I
   I > I   [ D$C$C$   ← $W
   I > I   [ D$C$     ← $W
6  I-> I   [ D$       ← $W
   I > E   ]          ← $W
   I > ^I  (          ←
   Y > I   (          ← #
   Y > ^I  ]          ← X
4  Y > ^I  (          ← L$C
2  Y > ^Y  (          ← R$C
3  Y > I   ]          ← $L$^I
2  Y > I   [ D$C$C$   ← $L+$W
3  Y > I   [ D$C$     ← $L+$W
2  Y > I   [ D$       ← $L+$W
2  Y > I   [ D$C$     ← $L$HE$B
4  Y > I   [ D$C$     ← $L$H$^I
6  Y > I   ]          ← $W
   Y > ^I  (          ←
   0 > 0   (          ← $B
   0 > ^A  ]          ← X
   0 > ^0  ]          ← LL#
   0 > ^A  ]          ← LL
   0 > 0   (          ← L$C
   0 > ^Y  ( W        ← R$C
   0 > ^0  ]          ← R
3  0 > 0   ]          ← $LI$W
   0 > 0   ]          ← $L+I$W
   0 > 0   ]          ← $LE$W
   0 > 0 ~ ]          ← $L+E$W
   0 > 0   ]          ← $L$^I
3  0 > ^A  [ D$C$C$   ← $L$^N
6  0 > ^A  [ D$C$     ←  $L$^N
   0 > 0   [ D$C$C$   ← $L+$W
   0 > 0   [ D$C$     ← $L+$W
   0 > 0   [ D$       ← $L+$W
2  0 > 0   ]          ← $EE$B
   0 > 0   [ D$C$     ← $L$W#
2  0 > 0   [ D$C$     ← $L$HE$B
3  0 > 0   [ D$C$     ← $L$H$^I
   0 > ^0  (          ← ^G
2  0 > 0   ]          ← $W
   0 > ^A  (          ←
   U > U   (          ← #
4  U > ^Y  ]          ← X
   U > J^U [ Z$       ← R$W
   U > ^U  [          ← R$W
   U > JU  [ Z$       ← $L$W
   U > U   [          ← $L$W
2  U > JUW [ Z$       ← $W
2  U > UW  [          ← $W
   U > 'U  [ G$       ← L$L
   U > ^Y  (          ←
   W > W   (          ←

   %PREFIX
   ANTI
5  BE
   BI
5  CIRCUM
   COM
   CON
   CO
   DE
   DIS
   EM
   EN
   EXTRA
   EX
4  HYPER
   IM
```

```
    INTER
4   INTRA
    IN
6   MAL
1   MAXI
5   MICRO
1   MINI
    MIS
2   MULTI
1   NEO
    NON
    PER
    PRE
    PRO
2   PSEUDO
5   PSYCHO
    RE
3   SELF
2   SEMI
4   SUB
3   SUPER
    TRANS
4   ULTRA
    UNDER
4   UNI
    UN

    %SUFFIX
    F 117-S'
    F 10-IFIC
    F 14-IC
    F 54-OID
    E 36-ED
1   F 50-FOLD
6   P 41-HOOD
    F 60-ACE
4   F 160-ICE
    F 41-ANCE
    F 45-IENCE
7   F 41-ENCE
2   F 20-ESCE
6   F 41-EE
    F 60-AGE
3   F 114-LIKE
    F 13-ABLE
    F 13-IBLE
5   F 40-PHILE
5   F 160-ILE
    F 41-ULE
3   P 10-SOME
    F 67-URE
```

```
    2   F 100-WISE
        F 77-ATE
    3   F 45-ETTE
    5   F 70-UTE
    2   F 14-ESQUE
    2   F 10-IFICATIVE
    5   F 57-ATIVE
        F 52-IVE
    4   F 20-OGIZE
    2   F 20-OPHIZE
    3   F 20-OLIZE
    2   F 20-OMIZE
    1   F 20-OQUIZE
        F 25-IZE
    3   F 100-SELF
        F 66-ING
        F 40-LOG
        F 15-ISH
    1   F 50-ETH
    5   F 10-EAL
        F 55-IAL
    5   F 10-UAL
        F 55-AL
    5   F 44-FUL
    1   F 40-DOM
        F 45-ISM
    4   F 44-ARIAN
        F 55-IAN
        F 50-AN
        E 30-EN
    5   F 70-EON
        F 47-IFICATION
        F 44-ITION
        F 75-ION
    5   F 40-OON
        F 44-SHIP
    3   F 50-ULAR
        F 50-AR
    4   F 64-EER
    4   F 40-OMETER
    2   F 40-METER
        E 53-ER
        E 42-OR
        F 134-'S
        F 166-ES
        F 141-OSIS
    6   F 140-SIS
    2   F 141-ITIS
    5   F 140-IS
    4   F 14-LESS
        P 41-NESS
```

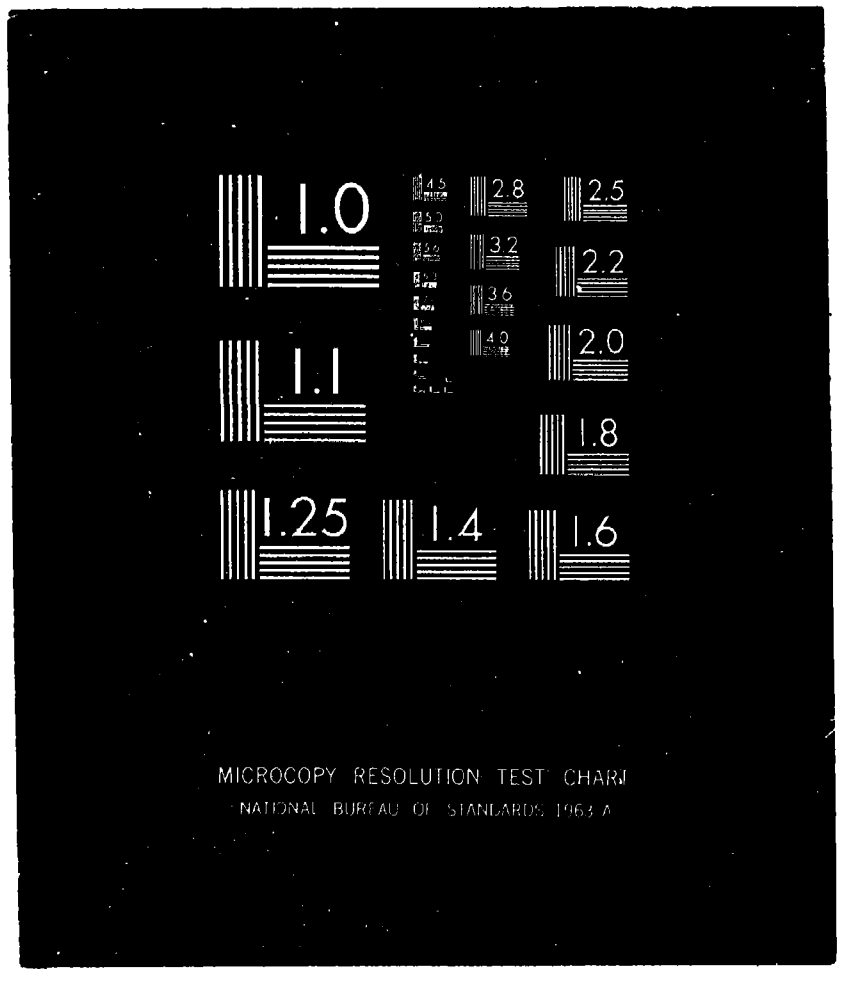
```
    5   F 10-EOUS
    7   F 15-IOUS
    3   F 15-UOUS
        F 17-OUS
    5   F 174-US
        F 167-S
    4   F 40-LET
        F 51-ANT
        F 55-IENT
        P 42-MENT
        F 51-ENT
        F 111-EST
        F 55-IST
    1   F 111-MOST
    3   F 47-CRACY
        F 50-EY
    5   F 21-IFY
    3   F 47-ARCHY
    3   F 10-WORTHY
    5   F 103-ABLY
    5   F 102-IBLY
        P 15-LY
    1   F 47-IMONY
        F 54-ARY
    5   F 54-ERY
    4   F 51-ATORY
        F 56-ORY
    5   F 40-ATRY
    2   F 40-OMETRY
    2   F 40-METRY
        F 44-RY
    3   F 47-IETY
        F 43-ABILITY
        F 43-IBILITY
        F 41-ITY
    1   P 47-PLASTY
        F 41-TY
    1   F 47-OQUY
        F 55-Y
```

AUTHOR INDEX - AJCL

AUTHORS OF SIGNED CONTRIBUTIONS

1974-1976 - Microfiches 1 - 58

END


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A